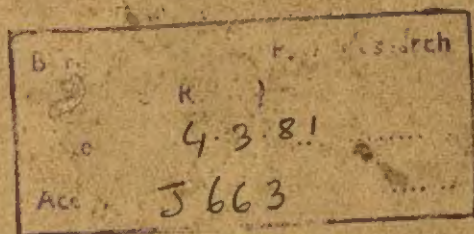
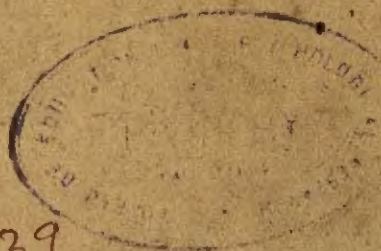


FEBRUARY 1969  
VOL. XXXIX, NO. 1

# Review of Educational Research



EDUCATION OF  
EXCEPTIONAL CHILDREN



Vol. 39  
1969

AMERICAN  
EDUCATIONAL RESEARCH ASSOCIATION

\$1000



## EDITORIAL BOARD

**Editor and Chairman:** JACOB T. HUNT, Professor of Education, University of Georgia, Athens, Georgia

RICHARD A. DERSHIMER, Executive Officer, American Educational Research Association, 1126 Sixteenth Street, N.W., Washington, D.C. 20036

JOHN I. GOODLAD, Dean, Graduate School of Education, University of California, Los Angeles, California

JACK V. EDLING, Professor, Teaching Research, Oregon State System of Higher Education, Monmouth, Oregon

JOHN W. COTTON, Professor of Education and Psychology, University of California, Santa Barbara, California

---

## INTERNATIONAL CONTRIBUTING EDITORS

MIECZYSLAW CHOYNOWSKI  
Polish Academy of  
Sciences  
Warsaw, Poland

JOHN DEAKIN  
West African  
Examinations Council  
Accra, Ghana

SIPKE D. FOKKEMA  
Free University  
Amsterdam, Holland

KJELL HARNQVIST  
Pedagogical Institute  
Goteberg, Sweden

TORSTEN HUSEN  
University of Stockholm  
Stockholm, Sweden

YUSEF SALAH EL-DIN  
KOTB  
Ein Shams University  
Cairo, United Arab  
Republic

M. J. LANGEVELD  
University of Utrecht  
Utrecht, Holland

ANTOINE LEON  
Sorbonne  
Paris, France

ALBERTO MARZI  
Institute of Psychology  
Florence, Italy

PRAYAG MEHTA  
National Council of  
Educational Research  
Delhi, India

B. F. NEL  
University of Pretoria  
Pretoria, South Africa

LORENZO B. PARADES  
University of Centro  
Escolar  
Manila, Philippines

EDWIN A. PEEL  
University of Birmingham  
Birmingham, England

JOHN PILLEY  
University of Edinburgh  
Edinburgh, Scotland

GENCHO D. PIRYOV  
Sofia University  
Sofia, Bulgaria

R. L. PLANCKE  
University of Ghent  
Ghent, Belgium

COLSELL SANDERS  
University of Western  
Australia  
Nedlands, Western  
Australia

JOHANNES SANDVEN  
University of Oslo  
Oslo, Norway

FRED W. SCHMID  
University of Zurich  
Zurich, Switzerland

WALTER SCHULTZE  
International High School  
Frankfurt, Germany

W. A. C. STEWART  
University of Keele  
Keele, Staffordshire,  
England

R. E. STROOBANT  
Victoria University  
Wellington, New Zealand

OLADELE TAIWO  
Federal Advanced  
Teachers College  
Yaba, Nigeria

HARALD TORPE  
Royal Danish College of  
Education  
Copenhagen, Denmark

PHILIP VERNON  
University of London  
London, England

GEERT WIELANGA  
Free University  
Amsterdam, Holland

---

Published five times a year by the American Educational Research Association,  
1126 Sixteenth Street, N.W., Washington, D.C. 20036.

Copyright © 1969 by the American Educational Research Association.

Second class postage paid at Washington, D.C.



# REVIEW OF EDUCATIONAL RESEARCH

*Official Publication of the American Educational Research Association*  
*Contents are listed in the Education Index.*

---

Volume XXXIX, No. 1

February 1969

---

## Education of Exceptional Children

Reviews the literature for the three-year period since the issuance of  
Vol. XXXVI, No. 1, February 1966

### TABLE OF CONTENTS

Chapter	Page
Foreword .....	4
I. The Mentally Retarded .....	5
HERBERT J. PREHM, <i>University of Oregon, Eugene, Oregon</i>	
JAMES E. CROSSON, <i>University of Oregon, Eugene, Oregon</i>	
II. The Gifted .....	25
EDWARD C. FRIERSON, <i>George Peabody College for Teachers,</i> <i>Nashville, Tennessee</i>	
III. The Speech Handicapped .....	38
CLARK STARR, <i>University of Minnesota, Minneapolis, Minnesota</i>	
IV. The Visually Handicapped .....	52
CARSON Y. NOLAN, <i>American Printing House for the Blind,</i> <i>Louisville, Kentucky</i>	
SAMUEL C. ASHCROFT, <i>George Peabody College for Teachers,</i> <i>Nashville, Tennessee</i>	

<i>Chapter</i>	<i>Page</i>
V. Learning Disabilities .....	71
CORRINE E. KASS, <i>University of Arizona, Tucson, Arizona</i>	
VI. Behavior Disorders .....	83
JOHN P. GLAVIN, <i>Temple University, Philadelphia, Pennsylvania</i>	
HERBERT C. QUAY, <i>Temple University, Philadelphia, Pennsylvania</i>	
VII. The Deaf and the Hard of Hearing .....	103
STEPHEN P. QUIGLEY, <i>University of Illinois, Urbana, Illinois</i>	
Index .....	124



# REVIEW OF EDUCATIONAL RESEARCH

Vol. 30, No. 2

April 1969

## GUIDANCE AND COUNSELING

### Foreword

1: Theoretical Foundations of Guidance .....	127
Martin R. Katz	
2: School Guidance Programs .....	141
H. B. Gelatt	
3: Higher Education Programs and Student Development .....	155
Harold A. Korn	
4: Counselor Education .....	173
John M. Whiteley	
5: Changes Through Counseling .....	189
Ray E. Hosford	
Alan S. Briskin	
6: Group Counseling .....	209
Alan R. Anderson	
7: Career Development .....	227
John L. Holland	
Douglas R. Whitney	
8: Counseling Students With Special Problems .....	239
D. David Island	
9: Systems Approaches in Guidance .....	251
William W. Cooley	
Raymond C. Hummel	
10: Relevance and Research in Counseling .....	263
Carl E. Thoresen	

# REVIEW OF EDUCATIONAL RESEARCH

Vol. 39, No. 3

June 1969

## CURRICULUM

### Foreword

- 1: History of Curriculum Thought and Practice .....283  
*Arno A. Bellack*
- 2: Forces Influencing Curriculum .....293  
*John D. McNeil*
- 3: Curriculum Materials .....319  
*W. James Popham*
- 4: Curriculum Evaluation .....339  
*Robert L. Baker*
- 5: Methodological Issues in Curriculum Research .....359  
*Richard E. Schutz*
- 6: Curriculum: State of the Field .....367  
*John I. Goodlad*



# SCIENCE AND MATHEMATICS EDUCATION

## FOREWORD

LEARNING STUDIES IN SCIENCE EDUCATION .....	377
<i>Maurice Belanger</i>	
CURRICULUM DEVELOPMENT AND INSTRUCTIONAL MATERIALS .....	397
<i>Herbert A. Smith</i>	
TEACHER EDUCATION IN SCIENCE .....	415
<i>Matthew H. Bruce</i>	
CURRICULUM EVALUATION .....	429
<i>Wayne W. Welch</i>	
FOREIGN AND INTERNATIONAL PROGRAMS .....	445
<i>R. Will Burnett</i>	
PHILOSOPHICAL AND HISTORICAL BASES OF SCIENCE TEACHING .....	459
<i>James T. Robinson</i>	
CURRENT RESEARCH IN MATHEMATICS EDUCATION .....	473
<i>Thomas A. Romberg</i>	
CONDITIONS OF LEARNING IN MATHEMATICS: SEQUENCE THEORY DEVELOPMENT .....	493
<i>Ralph T. Heimer</i>	
ACTIVITY LEARNING .....	509
<i>Thomas E. Kieren</i>	
PROBLEM SOLVING IN MATHEMATICS .....	523
<i>Jeremy Kilpatrick</i>	
CLASSROOM TEACHING OF MATHEMATICS .....	535
<i>James Fey</i>	

VOL. 39, NO. 5 DECEMBER 1969

Diary N 17  
Date 17 - 4 - 70  
File Lu  
Bureau Ednl. Res. Research

# METHODOLOGY OF EDUCATIONAL RESEARCH

## FOREWORD

RESEARCH METHODOLOGY—A POINT OF VIEW..... 553  
*D. Bob Gowin*  
*Jason Millman*

HISTORY OF EDUCATION ..... 561  
*Mark Beach*

AESTHETIC INQUIRY ..... 577  
*David W. Ecker*  
*Thomas J. Johnson*  
*Eugene F. Kaelin*

ANTHROPOLOGICAL APPROACHES TO THE  
STUDY OF EDUCATION ..... 593  
*Peter S. Sindell*

PHILOSOPHY, EDUCATION, AND THE DOOMSDAY THREAT ..... 607  
*Leonard Joseph Waks*

POLITICS OF EDUCATION ..... 623  
*Michael W. Kirst*  
*Edith K. Mosher*

ECONOMICS OF EDUCATION ..... 641  
*Mary Jean Bowman*

SOCIOLOGY OF EDUCATION ..... 671  
*John Harp*  
*Stephen Richer*

EDUCATION'S CHALLENGE TO PSYCHOLOGY: THE  
PREDICTION OF BEHAVIOR FROM  
PERSON-ENVIRONMENT INTERACTIONS ..... 695  
*James V. Mitchell*

STATISTICS ..... 723  
*Donald L. Meyer*  
*Maryellen McSweeney*  
*Leonard A. Marascuilo*  
*Raymond O. Collier, Jr.*  
*Robert C. Larson*  
*Maurice M. Tatsuoka*



This issue of the REVIEW was prepared by the Committee on  
Education of Exceptional Children

MAYNARD C. REYNOLDS, *Chairman*, University of Minnesota, Minneapolis,  
Minnesota

JACOB T. HUNT, University of Georgia, Athens, Georgia

JAMES J. MCCARTHY, University of Wisconsin, Madison, Wisconsin

CARSON Y. NOLAN, American Printing House for the Blind, Louisville,  
Kentucky

WILBERT PRONOVOST, Boston University, Boston, Massachusetts

HERBERT C. QUAY, Temple University, Philadelphia, Pennsylvania

with the assistance of

SAMUEL C. ASHCROFT, George Peabody College for Teachers, Nashville,  
Tennessee

JAMES E. CROSSON, University of Oregon, Eugene, Oregon

EDWARD C. FRIERSON, George Peabody College for Teachers, Nashville,  
Tennessee

JOHN P. GLAVIN, Temple University, Philadelphia, Pennsylvania

CORRINE E. KASS, University of Arizona, Tucson, Arizona

HERBERT J. PREHM, University of Oregon, Eugene, Oregon

STEPHEN P. QUIGLEY, University of Illinois, Urbana, Illinois

CLARK STARR, University of Minnesota, Minneapolis, Minnesota

## FOREWORD

This is a time of great ferment in Special Education, both in practical program operations and in research. During the past two decades, programs for handicapped and gifted children have doubled and even tripled their size in many states. The U.S. Office of Education, many state departments of education, and several hundreds of colleges and universities have been strengthened in their Special Education departments by infusions of federal money. The Division of Research of the Bureau of Education for the Handicapped of the U.S. Office of Education is now the major source of support for research in the field. The history of this research program traces back more than a decade.

It is time now to reflect critically on what this surge of development within the new framework of strong federal leadership and supports amounts to and to reassess present trends.

It appears quite clear, even now, that organization for research efforts should and will be broadened in the form of special research centers and programs, with relatively less support for single projects. Implicit in this trend is the need to delineate more carefully the truly major problem areas and to mobilize multifaceted research around such areas. In practical program operations also, major changes are in the wind, with more shift to an explicit educational emphasis in all phases of diagnosis, classification, and instruction for exceptional children. This might well be nearly the last review on exceptional children whose chapter titles carry the traditional category labels borrowed from other fields. But the clearest prediction is that the next review in Special Education will be an interesting story!

MAYNARD C. REYNOLDS, *Chairman*  
*Committee on Education of Exceptional*  
*Children*



# CHAPTER I

## The Mentally Retarded

HERBERT J. PREHM and JAMES E. CROSSON\*

Previous issues of this REVIEW have consistently noted that the preceding three-year period had witnessed an acceleration in research activity on the mentally retarded. The rate of acceleration is increasing geometrically. Undoubtedly, no previous period has seen such an unbelievable volume of published research reports. Because of the volume of research reported, this review focuses almost exclusively on research reports published in either journal, monograph, or book form. It was not possible to review unpublished project reports, but Guskin and Spicker (1968) have recently reviewed the more relevant of these reports. For a more detailed consideration of some of the topics included in this chapter, the reader is referred to an excellent series of reviews edited by Ellis (1966, 1968).

### Classification

#### The Nature of Mental Retardation

Differing conceptualizations of the nature of mental retardation continue to underlie research pertaining to the classification of the retarded. Skeels (1966) has recently added a major supporting element to the "nurture edifice" which has, as its cornerstone, his contributions of the past twenty years. This follow-up study shows that divergent patterns of competencies, the result of enriched versus deprived developmental stimulation, were maintained in the experimental and control groups nearly 30 years after the experimental intervention.

In contrast to traditional approaches to the problem of defining mental retardation, Clausen (1966) advocated the analysis of ability structures as they are manifested in differential constellations of impaired and intact functions. Using a battery of 31 tests ranging from basic sensory and motor tasks through reasoning, memory, and integrative functions, he presented a comprehensive mapping of psychological functions which underlie intellectual development.

Another highly provocative alternative to the more traditional conceptions of mental retardation was offered by Bijou (1968) who contended that mental retardation should be regarded "not as a symptom, but as a form of behavior." The quality and quantity of opportunities for contact

\*Preparation of this chapter was facilitated by Research and Training Center Grant RT-16 from the Social and Rehabilitation Service of the U.S. Department of Health, Education, and Welfare, Washington, D.C. 20201.

were seen as the determinants of a person's mode of development. Factors which contribute to delays and failures in development, according to Bijou, are abnormal anatomical structure and physiological functioning, insufficient reinforcement and discrimination histories, disadventitious reinforcement of undesirable behaviors, and severe aversive stimulation.

### Individual Intelligence Tests

A considerable body of research literature pertaining to the validity, reliability, and diagnostic utility of various standardized intelligence scales with the mentally retarded has been published within the preceding years. Davis (1966) analyzed internal consistency of subtests of the *Wechsler Intelligence Scale for Children (WISC)*, as well as the total Verbal, Performance, and Full Scale scores of a group of 142 subjects (Ss). The measures were generally found to possess satisfactory reliability across three functional levels (borderline, mild, and moderate mentally retarded). Himelstein (1968) has concluded that the relationship between the *Stanford-Binet* (Form L-M) and scholastic achievement, social maturity measures, and developmental schedules is substantial, suggesting the instrument adequately predicts learning and other facets of intelligence in the mentally retarded. His review also indicated that it was difficult to evaluate the instrument's ability to discriminate at very low levels of mental ability.

The relationship between IQ's obtained on the *Peabody Picture Vocabulary Test (PPVT)* and those obtained on the *WISC* or the *Stanford-Binet* were examined by Shaw, Matthews, and Klove (1966), Rice and Brown (1967), and Brown and Rice (1967). The three studies led to the conclusion that the *PPVT* is not a valid test of intelligence as measured by the standard individual intelligence scales. However, Koh and Madow (1967) obtained an  $r$  of 0.93 between *PPVT* and *Stanford-Binet* mental age (MA), with a partial  $r$  of 0.90 when chronological age (CA) was held constant. The MA relationship was observed to be slightly curvilinear, with *PPVT* underestimating the *Stanford-Binet* below the five-year level and overestimating above nine and one half years. Hammill and Irwin (1967) found that the *PPVT* and *Stanford-Binet* were equivalent for both trainable mentally retarded (TRM) and educable mentally retarded (EMR) Ss, with the exception of educable Ss with diagnosed (or inferred) brain involvements over 10 years of age, and that *PPVT* rawscore test-retest reliability was acceptable over one-, two-, and three-year periods.

### Short-Form Intelligence Scales

Cole and others (1967) employed a sample of 165 elementary school children, being evaluated for possible special class placement, to determine which of the *WISC* subtests correlated most highly with full-scale IQ. Results were consistent with data reported in the *WISC* manual (i.e., the



Information, Comprehension, and Vocabulary scores yielded the highest correlations, with the exception that Picture Arrangements, rather than Arithmetic, correlated most highly in the present study). Silverstein (1968) employed a different method of shortening the *Wechsler Adult Intelligence Scale (WAIS)*, based on a split-half abbreviation, wherein the odd-numbered items of each subtest were retained. The resulting short-form test IQ's of 50 institutionalized Ss closely approximated the IQ's obtained on the entire instrument. Intercorrelations between the abbreviated and standard instruments, including the initial testing and a retest nearly four years later, were acceptably high.

### Sociometric and Personality Inventories

A revision of the *Syracuse Scales of Social Relations* for use with the mentally retarded was described by De Jung (1966). The instrument was reported to offer several advantages over the more traditional sociometric nomination techniques, including the availability of interrater and intergroup comparisons of both ratings made of others and ratings received from others. Analyses of data obtained by use of the modified instrument with third- and fourth-grade children and special class retardates led to the conclusion that the instrument is useful down to the third-grade level and for all but a few of the educable retarded pupils if it is individually administered.

Zigler, Butterfield, and Goff (1966) reported on a series of studies evaluating a scale designed to rate social deprivation. In the first study, a factor analysis of 80 preadmission institutional social histories was subjected to the Varimax rotation. Factors emerging were overt early rejection, later rejection, paternal competence, and maternal competence. A cross-validation based on samples drawn from two additional institutions yielded similar factor groupings, with the exception that the fourth factor was identified in terms of marital relationships of the parents and the parents' attitude toward the child.

Gardner (1967) questioned the use of the *California Test of Personality* with mentally retarded persons on the basis of their limitations in language skill and self-perception, as well as experiential differences in environmental background. Gozali and Bialer (1968) employed an item-reversal technique to assess the independence of Locus of Control Scale scores from the possible tendency to agree indiscriminantly among a mentally retarded sample. Results demonstrated a probable absence of social desirability effects on Locus of Control responses.

### Nonstandardized Instruments and Direct Measurement

A direct measure of work performance based on the number of production units minus task-related errors was employed as the criterion variable

in samples of mildly retarded Ss (Elkin, 1968) and moderately and severely retarded Ss (Elkin, 1967) to determine the value of the work-sample technique, as well as a large number of objective measures, in predicting productivity. For both samples, intellectual, psychomotor, and work-sample measures were the best predictors. However, when eight of the best prediction variables were analyzed with respect to their relationship to work-evaluation ratings obtained in actual community work placements, only the *O'Connor Finger Dexterity Test* was significantly related.

Several writers have reported on observation procedures based on techniques of functional analysis. While the procedures are unified with respect to behavior principles and associated technology, some variation in procedures is noted—principally in relation to the particular targets or objectives of analysis. Peterson (1967) described an observation and recording technique designed to assist nurses in planning, evaluating, and refining nurse-patient interactions. Hamerlynck, Martin, and Rolland (1968) reported on the usefulness of time-sampling measures of molar classes of incompatible behaviors (task-oriented and non-task-oriented) in a special education classroom. Crosson (1968a) provided a general discussion of functional analysis techniques and procedures for use by special education teachers and supervisors.

A new dimension of measurement of the classifications of the mentally retarded was described by Leland and others (1967). This dimension—adaptive behavior—is the subject of a major nationwide research project jointly sponsored by the AAMD and NIMH, an outgrowth of the work of the AAMD Project on Technical Planning in Mental Retardation. A technical description of the recently completed experimental version of the Adaptive Behavior Scales is given by Nihira, Foster, and Spencer (1968).

## Education and Rehabilitation

### Placement

The disappointing lack of meaningful research results bearing on the efficacy of special class placement has apparently led to the emergence of new trends in this line of investigative activity. One trend is to re-examine research strategies used to determine efficacy. Meyers and Meyers (1967) focused on the problems of divergent conceptions of the educable mentally retarded as they contribute to problems in school placement and suggested a specific deficit (or ability) approach to classification as an alternative to contemporary placement criteria. A second trend is reflected in the investigation of unitary, rather than generic, experimental variables. Mayer (1966) studied the relationship of early special class placement and self-concept with disappointing results: the hypothesis that early placement in special classes develops more positive self-concepts was not supported.



## Curricular Adaptations

Several writers have investigated psychological processes in the context of curricular areas. Sheperd (1967) employed an extensive battery of tests to assess selected factors in the reading ability of EMR boys. He compared two groups of 20 Ss matched on relevant criteria, who differed in their reading skill. Measures of silent and oral reading, articulation, sound-blending ability, ability to use context clues, fund of basic information, and social and emotional adjustment significantly favored the "adequate" group. The two groups did not differ in performance on the Memory-for-Designs Test, visual closure, handedness, eyedness, or lateral dominance. Performance of both groups on tests of automatic-sequential psycholinguistic functions was below their mean MA level.

Edmonson and others (1967) reported on the development of a test of social inference and a program of training in social cue interpretations for the retarded. The authors reasoned that the retardate's characteristic insensitivity to social cues mediates against his admission to social transactions. They have determined that this deficit can be assessed by means of the retardate's interpretive verbal response to pictures depicting interpersonal situations. A curricular unit designed to illustrate the cultural modes of information exchange via signals resulted in gains in social-signal decoding by experimental Ss which exceeded the gains made by contrast groups. Although clear-cut evidence of social behavior decoding gains did not result, experimental Ss received higher post-treatment ratings of social behavior more frequently than contrast Ss.

The use of programed instruction with the mentally retarded has also received considerable attention within the past three years. Bijou and others (1966) reported on a three-year project on programed reading, writing, and arithmetic instruction for a group of mentally retarded children. The paper describes the development of an effective motivational system, programed procedures for establishing appropriate study behaviors, and instructional materials. Johnson (1966) investigated the relative effects of programed instruction with and without additional teaching by lesson plans in comparison to regular instruction in addition and subtraction. He concluded that the use of programed instruction techniques in conjunction with additional teaching may be more effective with the educable mentally retarded than the use of conventional approaches.

Hewett, Mayhew, and Rabb (1967) described an experimental reading program that had been used with several types of educationally handicapped children, including retardates. All Ss showed steady progress toward the acquisition of a 155-word vocabulary. Several alterations of reinforcement conditions were performed to assess their effects on acquisition. Removal of primary reinforcement (candy) and response cost (involving the taking away of candy) had little effect on performance. On the other hand, a "time hold" contingency produced an increased response rate. Fangman (1967)

used tachistoscopically presented, meaningful word phrases in a study of the effects of phrase-reading training on reading speed, accuracy, and comprehension. The author concluded that retardates increase their reading rate significantly, with no loss in comprehension, by means of automated phrase-reading training. He found that increased reading rate produced by the method is not associated with intelligence measures.

Special-purpose curricular innovations were evaluated by several writers. Corter and McKinney (1968) developed a flexibility training program designed to improve cognitive flexibility through increased ability to shift concepts. The training program apparently was effective in producing increases in cognitive flexibility. A hypothesized difference in response to training between retarded and nonretarded Ss did not occur, and the effects of training appeared to generalize to other abilities. Methodological limitations prevented clear-cut interpretations of these effects.

In the area of vocational training, Knaus (1967) compared three methods of presenting programmed vocational instruction to adolescent MR females, concluding that the automated method is most efficient when achievement speed is critical, but that an integrated approach wherein the automated program is used as an adjunct to partially programmed lectures, discussions, and demonstrations is most efficient in producing strong performance on a work-sample test. Vocational prognosis and adjustment were explored by Bae (1968), who compared instructor's ratings with indirect measures of vocational adjustment. Bae's study led to the conclusion that factors associated with vocational efficiency vary among training programs and points to the implicit conclusion that emphasis should be placed on differential abilities rather than global vocational efficiency as relevant measures of vocational adjustment.

Switzer (1966) discussed the coordination of vocational rehabilitation and special education services for the mentally retarded. Bloom (1967) evaluated one such program, concluding that the work-study approach had been generally effective. A finding of possible interest to others planning such programs is that students who were paid for work accomplished while in job training received better ratings from their supervisors on measures of responsibility and emotional stability than those trained without payment.

## **Behavior Modification**

Behavior modification, a technological outgrowth of the science of operant behavior, has continued in the trend of increasing application in the field of mental retardation (see Doubros and Daniels, 1966; Giles and Wolf, 1966; Karen and Maxwell, 1967; Minge and Ball, 1967). There also appears to be a trend toward a broadening of the range of behavioral targets and the complexity of application. Whitney (1966) has discussed the role of behavior technology in nursing. Johnson, Haughton, and LaFave (1965), Huddle (1967), and Crosson (1968b) have described the appli-



cation of operant principles in sheltered workshops. Perhaps the most significant recent applications are those which attempt total environmental programming. Lent (1968) described one such program—perhaps the first—designed for a cottage of 71 TMR female residents. A graduated token economy was employed which was integrated at more advanced stages of programming with the normal reinforcement systems operating in the extra-institutional community.

### Postschool Adjustment

Heading the list of research investigating the postschool adjustment is the latest addition to the classic study in the area: Baller, Charles, and Miller (1967). Available Ss from original "low," "middle," and "high" samples drawn in 1934 were assessed with respect to their mid-life attainment. The study showed that about 65 percent of the low group had remained in or returned to their home community, while only 40 percent of the other two groups had remained. Nine of the available Ss were institutionalized. Nearly a third of located low-group Ss were deceased, but the mortality rate was low for the middle and high groups. The divorce rate was two to ten times above the national average. While unemployment and welfare subscription were somewhat above the national average for the low group, the subjects had generally shown steady employment over the years. All Ss were generally law-abiding. It was further noted that the more successful of the low group Ss were likely to have acquired a skill early and worked at it continually and to have remained in one community.

An interesting study by Chaffin (1967) demonstrated a positive relationship between production rate and employer's judgments of success or nonsuccess. Behavior modification procedures were employed to reverse the production rates of the Ss after they had been judged by their employers. The results indicated that an employee who had been judged as successful and then exhibited a lower production rate was subsequently judged as nonsuccessful by the employer. A converse trend was noted for Ss initially judged unsuccessful.

## Psychological Processes

### Learning

*Motor Performance.* A major concern of the special educator involves the establishment of conditions which will facilitate the motor development and performance of the mentally retarded. Lillie (1968) administered a series of motor-development lessons to a group of 16 culturally disadvantaged, preschool-age, mentally retarded youngsters. Pre- and post-treatment *Lincoln-Oseretsky Motor Development Scale* scores of preschool-age retarded Ss, a regular kindergarten, and a home control group were compared. The post-test gross motor performance of all three groups were

comparable. The three groups did differ significantly on post-treatment fine motor performance. The specially designed training materials resulted in the retarded Ss having fine motor coordination which was superior to either control group.

In comparison to other areas of motor performance, reaction time (RT) has been the subject of considerable investigation during the past three years. Berkson and Baumeister (1967) and Baumeister and Kellas (1968) have shown that the RT performance of the retarded is not only slower but is also more variable than that of Ss of normal intelligence. Several studies have investigated the relationships between RT performance and warning signal durations or warning intervals. In a series of two experiments using institutionalized male retardates, Kellas and Baumeister (1968) obtained data which indicated that regular RT procedures provide faster RT's at short warning intervals and that as the warning interval increases there is also an increase in RT. In contrast, under randomized warning interval durations, the shortest (0 second) and longest (8 seconds) preparatory intervals produced the slowest reaction time. Baumeister, Dugas, and Erdberg (1967) have also shown that increases in warning signal intensity result in longer reaction times.

*Perception.* Terdal (1967) has found that although nonretarded Ss show a more marked preference (as defined by duration of looking time) for complex designs, both retarded and nonretarded Ss prefer complex to simple stimuli. Winters and Gerjuoy (1967) compared left-to-right eye movement and verbal responses of retarded adolescent Ss with those of equal-CA and -MA groups of normal Ss. The task involved the tachistoscopic presentation of alphabet and nonalphabet stimuli. Their data indicated that the normal Ss eye movements and verbal reports were from left to right significantly more often than those of their retarded Ss. They also found that retardates who were more organized in their eye movements and verbal reports also made more correct responses than did disorganized retardates. Disorganized retardates also scored lower on the reading subtest of the *Wide Range Achievement Test*. Spitz (1967) has found that equal-CA retarded and normal Ss are equally susceptible to the distorted-room illusion; and Wachs and Cromwell (1966) have found that subsequent to a failure experience, retardates exhibited significantly more perceptual distortion than did normal Ss.

*Rote Verbal Learning.* McManis (1966) and Sen, Clarke, and Cooper (1968) have investigated the effects of isolating items in lists of verbal material learned using the serial anticipation method. McManis found that isolation of a list item for meaningfulness, color, or both produced a significant isolation effect for both retarded and normal Ss. Sen, Clarke, and Cooper found that a serial list comprised of two color and six black-and-white pictures was learned at a significantly more rapid rate than were lists comprised of all black-and-white pictures, or lists in which the isolated items were pictures outlined in red. They also found that isolation resulted



in fewer errors for that item. Semmel and Williams (1968) found that retarded Ss required a significantly greater number of trials to reach criterion on a serial learning task than did groups of CA-matched borderline and normal Ss.

Using a paired-associate learning task, Baumeister, Hawkins, and Davis (1966) found that while varied exposure durations did not affect the performance of normal Ss, exposure durations of 5-7 seconds facilitated the performance of the retardates to a greater extent than did exposure durations of 1-4 seconds. They also found that the performance of the retarded Ss was inferior to the performance of both CA- and MA-contrast Ss of normal intelligence.

Baumeister and Berry (1968) have found that, although a variety of stimulus conditions did not affect the learning performance of mild and borderline retardates, they did affect the performance of normal Ss. They also found that normal Ss generally performed significantly better than did the retarded Ss. Prehm (1967) has found that while type of list (picture-picture, picture-word, word-word) did not affect the learning performance of either retarded or normal Ss, spelling the response was more difficult for the Ss than pronouncing it as a word. In an earlier study, Prehm (1966a) found (a) that the paired-associate learning performance of normal Ss was significantly superior to that of retarded Ss; (b) that, at his high level of difficulty, normal Ss learned meaningful paired associates in fewer trials than nonmeaningful pairs; and (c) that meaningfulness exerted no statistically significant effect for retarded Ss.

Borkowski and Johnson (1968) showed that when mediators were not made available to retarded Ss, their paired-associate learning performance was inferior to that of normal Ss. Using a similar design, Penney, Seim, and Peters (1968) obtained results indicating that retardate mediation performance was inferior to that of normal Ss under a short S-R interval and superior under a long (12-second) interval.

*Discrimination.* The discrimination learning performance of the mentally retarded continues to be vigorously researched. Of particular interest is the performance of retardates on reversal or partial reversal shifts. Heal (1966) found intradimensional shifts to be significantly easier to make than extradimensional shifts. He also found that overtraining on the original problem aided the discrimination shift performance of nonretarded Ss but impaired the performance of retardates. Ohlrich and Ross (1966) found that overtraining on their original discrimination problem resulted in a reversal shift being easier to learn than a nonreversal shift. Heal, Ross, and Sanders (1966) studied the ability of retarded and nonretarded Ss to make partial (in which S approached or ignored a novel cue) and standard discrimination reversals. They found that retardates were significantly inferior to matched nonretarded Ss in overall reversal, that a reversal shift was more difficult than partial reversal shifts, and that the introduction of novel cues into the reversal tended to disrupt performance.

Schoelkopf and Orlando (1966) investigated the effects of delay of reinforcement on the performance of institutionalized retardates. The four delay periods (3, 6, 12, or 18 seconds) investigated did not affect performance in a linear fashion. Performance was poorest in the 3- and 18-second delay periods and best in the 6- and 12-second conditions.

*Learning Set and Transfer of Training.* Interest in the learning set performance of the retardate seems to have declined during the past three years. An explanation of the decline in interest is difficult to formulate, particularly in view of the vigorous activity in this area before 1966 (Kaufman and Prehm, 1966). Harter (1967) investigated the role played by MA, IQ, and motivation in discrimination learning-set performance. She found that the higher the IQ and the higher the MA, the faster a learning set was established. Although IQ and MA were the most important of the variables, she also found that providing social reinforcement to retarded Ss seemed to act as a distracting stimulus; the performance of retarded Ss in a standard learning-set condition was superior to that of retardates receiving social reinforcement.

In spite of the seeming importance of knowledge about the conditions under which retardates transfer learning, this topic of transfer remains one of the most underresearched in the area of mental retardation. Clarke and Cooper (1966) reported that when adjustments are made in task difficulty, retarded adolescents and adults demonstrate levels of transfer performance that are comparable to those shown by retarded children. Butterfield and Butterfield (1968) found that delay intervals (10 minutes, 24 hours) between training and testing did not affect transposition performance and that retarded Ss with MA's between 6-5 and 7-5 years were more able to verbalize the solution to a transposition problem than subjects with lower MA.

*Concept Learning and Problem Solving.* Until recently it was assumed that retardates were unable to manifest creative thought. However, research has shown that this assumption is erroneous. Cawley and Chase (1967) found that special class retardates, regular class retardates, and nonretarded Ss matched on MA exhibited comparable performance on a battery of productive thinking tests. Smith (1967) found that while regular class retardates scored on 22 Guilford-type tests of creative thought, their performance was significantly below that of CA-matched nonretarded Ss on 13 of the measures. With CA, IQ, level of academic achievement, and MA controlled through the use of the analysis of covariance, the retarded and nonretarded Ss differed only on word fluency.

Stephens (1966a, b) investigated the abilities of retarded and nonretarded Ss to identify which of several pictured items belonged to the same conceptual category. On an unstructured test (S was not told name of category) and a structured test, the performance of the retarded was significantly inferior to that of normal Ss matched for CA. The performance of retardates was comparable to that of MA-matched normals on two of



three conceptual categories for both types of tasks. In the second study, he found that the performance of retarded males was inferior to that of CA-matched normal Ss and that the retarded were unable to provide names for categories that they had previously used correctly.

Milgram (1966b) has found that while the picture-sorting performance of trainable-level retardates was comparable to that of educable and non-retarded Ss, their ability to verbalize the way in which the pictures were related was significantly inferior to that of the other two groups. Milgram (1966a) also found that altering the instructions given on a test of verbal conceptual ability did not affect performance on the test for either non-retarded or retarded Ss. Milgram and Furth (1967) found that while the ability to choose correctly two items representing a part-whole concept was comparable for retarded and nonretarded Ss at MA-6, retarded Ss were significantly inferior to nonretarded Ss at MA-9. Training Ss to frame the concept in a sentence facilitated the performance of normal Ss only.

*Operant Conditioning.* An appreciable number of recent studies have incorporated extensions of operant behavior principles to the experimental analysis of retarded behavior. Most studies assessed variables associated with reinforcement and their correlated effects on behavior. Watson, Orser, and Sanders (1968) assessed the reinforcement preference of severely retarded children and concluded that poker chips acquired the properties of a conditioned reinforcer when appropriately associated with candy or toy rewards. Steinman (1968) employed an alternating multiple-variable ratio schedule in an attempt to assess the relative strength of candy and verbal approval as reinforcers. After stable response rates were established, verbal reinforcement was presented as a discriminative stimulus for material reinforcement which was gradually eliminated. A post-treatment run on the multiple schedule failed to show an increase in the probability of responding in the verbal-approval component of schedules. The author concluded that the history of the natural acquisition of social reinforcers may be long and complex. Wiesen and others (1967) shaped a group of severely retarded Ss to present each other with M&M candies, their act of reinforcement being rewarded by the experimenter. A 5-day extinction period and 4 days of reconditioning followed 14 days of conditioning. Pairs of Ss, assigned on the basis of low initial rates of interaction, increased their rates of social responding as a function of the number of conditioning days.

Baldwin (1966) analyzed the effects of three types of reinforcement programs on the acquisition of social skills (as measured by the Progress Assessment Chart) in institutionalized retardates. Three reinforcement systems—a token economy, social reinforcement, and primary food reinforcement—were compared with a control condition in which Ss were maintained in their regular cottage programs. Although tangible reinforcers were generally most effective, a hierarchal effect was noted. Nutritive reinforcers were more effective for younger Ss, while tokens were more powerful with older Ss.

Gates and Fixsen (1968) used an array of eight response manipulanda to establish preferential response patterns in a sample of eight institutionalized retardates, then evaluated response variability as a function of different schedules of reinforcement. In general, it was concluded that responding is less variable under variable-interval reinforcement (the only intermittent schedule employed) than under continuous reinforcement. A general trend for decreased variability of responding as a function of length of experimental involvement was noted. Spradlin, Girardeau, and Hom (1966) attempted to analyze the stimulus function of a prior reinforcer during extinction. Stimulus function of the reinforcer was most clearly in evidence when delivered early in extinction. However, permanence of the effect was not demonstrated for all Ss.

Crosson (1967) employed principles associated with operant chaining to analyze and program vocational behaviors. Accuracy data revealed rapid acquisition of two complex tasks. Accuracy was assumed to be more strongly associated with discriminative efficiency of the task stimuli than reinforcement spacing. A second experiment revealed no significant difference with respect to maintenance of performance under social and extrinsic (token) reinforcement, and task conditions were assumed to be a function of schedule effects rather than characteristics intrinsic to the tasks.

## Memory

During the past three years, evidence has accumulated both for and against the hypothesis that the retardate exhibits a stimulus trace deficit (Ellis, 1963) which results in poor short-term memory. Fagan (1968) found that the immediate recall performance of educable retarded Ss was inferior to that of MA-matched nonretarded Ss and showed similar slopes on item-recall measurements and that the retarded Ss exhibited superior performance on ordered recall retention. In contrast, Belmont (1967) found that the perceptual short-term memory of institutionalized retardates did not differ from the performance of normal children or college students except at a two-second delay where the retarded were superior. Ellis and Anders (1968) found evidence suggesting that an encoding deficit may be responsible for observed retardate short-term memory deficit. Hawkins (1966) found that a "trace" score derived from *Wechsler* subtest items correlated with the recognition performance of brain-damaged retardates but not the performance of cultural-familial retardates.

Evidence supporting the hypothesis that retardates exhibit a long-term retention deficit was provided by Vergason (1966) and Prehm (1966b). Vergason found that the 24-hour and 30-day recall performance of educable retardates was significantly inferior to that of CA-matched nonretarded Ss. He also found that 400 percent overlearning facilitated the recall performance of both retarded and nonretarded Ss. Prehm found that the

relearning performance of educable retardates was significantly inferior to that of nonretarded Ss at both 24 hours and 3 months. He found that 100 percent overlearning did not result in improved relearning performance. Milgram (1967) found that while retarded and nonretarded Ss benefited from training in using a sentence to join the stimulus-response components of a paired-associate task, only nonretarded Ss retained the mediation set for a period of a week.

## Language and Communication

Research on the language abilities of the retarded has steadily increased during the past three years, particularly research on aspects other than psycholinguistic performance. Lovell and Bradbury (1967) found that the performance of adolescent retardates on Berko's test of English morphology was inferior to that of Berko's sample of first-grade pupils. Cartwright (1968) found that the written language of adolescent retardates was inferior to that of CA-matched nonretarded Ss in sentence length, ratio of type of word to number of words, number of adjectives and adverbs, and grammatical and spelling correctness. Nonretarded Ss matched with the retarded on MA were superior only on the type-number of words ratio and grammatical and spelling correctness. Semmel and others (1968) found that, while adolescent nonretarded Ss gave significantly more associations to verbal stimuli of the same form class as the stimulus than did a group of adolescent retardates, noninstitutionalized retarded Ss and MA-matched nonretarded Ss were similar in their performance.

## Personality

Much of the research reported in this area during the past three years confirms and extends somewhat the findings of earlier studies. Blount and Pielstick (1967) reconfirmed the findings that retardates exhibit better motor performance under a condition of high incentive than under low incentive. Siegel and Forman (1967) found that retardates, rewarded for correct responses on a two-choice discrimination task with either a penny or a washer, tended to fixate on the 50 percent chance of receiving the penny on each trial. Gardner and Brandl (1967) found that providing retardates with supportive comments as opposed to tangible rewards or instructions resulted in superior intentional and incidental learning. Huddle (1967) found that trainable retarded Ss who were given monetary reinforcement performed an industrial task at a significantly higher level than did Ss receiving no reward. McManis (1967) found that persistence on a marble-sorting task varied as a function of combinations of reproof, praise, and competition incentives. Pairs of institutionalized retarded Ss were most persistent when praise was paired with competition and least persistent when reproof was paired with praise.



Gardner (1968), in a scathing review and critique of research focused on the personality characteristics of the retarded, concluded that hard data bearing on the rate of occurrence of personal adjustment problems in retardates are virtually nonexistent. Although this conclusion may be questioned, the last year has provided neither the practitioner nor the researcher with data pertaining to this question. The reader is referred to Gardner's paper for a more extensive discussion of this literature.

### Social Adjustment

Meyerowitz (1967) has found that second-grade EMR children seem to be neither rejected nor sought as playmates within their own neighborhoods. He also found that regular class retardates were better known in their neighborhoods than special class retardates. Weiss and Weinstein (1968) found that both institutionalized and noninstitutionalized retardates attempt to manipulate the behavior of both peers and authority figures by the tactic of asking as opposed to aggression or proposing a situation where the other person's behavior would be reciprocated. Chennault (1967) found that by pairing unpopular special class retardates with popular members of their class in special organized group activities, the peer acceptance and perceived peer acceptance of the unpopular Ss improved significantly.

Levine, Elzey, and Paulson (1966) studied the *Cain-Levine Social Competency Scale* ratings of trainable retardates who were either successful or unsuccessful in public school programs. They found that the rated performance of successful TMR Ss was significantly higher than the rated performance of unsuccessful retardates and suggested that the scale could be profitably used as a primary criterion for determining the TMR's eligibility for school programs.

### Concluding Comments

The volume of research reported during the past three years attests to the viability of the area. While the body of knowledge about the retardate is growing, this same growth threatens our ability to keep abreast of the literature. The inevitable continuation of the current rate of production increases the importance of periodic literature reviews.

The level of sophistication of research questions has definitely increased. Sterile areas of research (e.g., the question of whether or not retardates learn at the same rate as do nonretarded persons) are being abandoned in favor of more promising lines of research. The result has been the development of a "little d," as opposed to a "large D," approach to the description of retardate deficits. In addition, researchers are becoming considerably more attentive to the validity of their research procedures and tools. Continuation of these trends is important if the quality of our knowledge is to be increased.

## Bibliography

- BAE, AGNES Y. "Factors Influencing Vocational Efficiency of Institutionalized Retardates in Different Training Programs." *American Journal of Mental Deficiency* 72: 871-74; May 1968.
- BALDWIN, VICTOR L. *Development of Social Skills in Retardates as a Function of Three Types of Reinforcement Programs*. Doctor's thesis. Eugene: University of Oregon, 1966. 77 pp. Abstract: *Dissertation Abstracts* 27: 2865A; No. 9, 1967.
- BALLER, WARREN R.; CHARLES, DON C.; and MILLER, EBERT L. "Mid-Life Attainment of the Mentally Retarded: A Longitudinal Study." *Genetic Psychology Monographs* 75: 235-329; May 1967.
- BAUMEISTER, ALFRED A., and BERRY, FRANKLIN M. "Context Stimuli in Verbal Paired-Associate Learning by Normal Children and Retardates." *The Psychological Record* 18: 185-90; April 1968.
- BAUMEISTER, ALFRED A., and KELLAS, GEORGE. "Distribution of Reaction Times of Retardates and Normals." *American Journal of Mental Deficiency* 72: 715-18; March 1968.
- BAUMEISTER, ALFRED A.; DUGAS, JEAN; and ERDBERG, PHILIP. "Effects of Warning Signal Intensity, Reaction Signal Intensity, Preparatory Interval and Temporal Uncertainty on Reaction Times of Mental Defectives." *Psychological Record* 17: 503-507; October 1967.
- BAUMEISTER, ALFRED A.; HAWKINS, WILLIAM F.; and DAVIS, PATRICIA A. "Stimulus-Response Durations in Paired-Associate Learning of Normals and Retardates." *American Journal of Mental Deficiency* 70: 580-84; January 1966.
- BELMONT, JOHN M. "Perceptual Short-Term Memory in Children, Retardates and Adults." *Journal of Experimental Child Psychology* 5: 114-22; March 1967.
- BERKSON, GERSHON, and BAUMEISTER, ALFRED. "Reaction Time Variability of Mental Defectives and Normals." *American Journal of Mental Deficiency* 72: 262-66; September 1967.
- BIJOU, SIDNEY W. "The Mentally Retarded Child." *Psychology Today* 2: 46-51; June 1968.
- BIJOU, SIDNEY W., and OTHERS. "Programmed Instruction as an Approach to the Teaching of Reading, Writing, and Arithmetic to Retarded Children." *Psychological Record* 16: 505-22; October 1966.
- BLOOM, WALLACE. "Effectiveness of a Cooperative Special Education Vocational Rehabilitation Program." *American Journal of Mental Deficiency* 72: 393-403; November 1967.
- BLOUNT, WILLIAM R., and PIELSTICK, N. L. "Magnitude of Reward and Reward Value in Male Retardates." *Perceptual and Motor Skills* 24: 931-42; June 1967.
- BORKOWSKI, JOHN G., and JOHNSON, LEE ODESS. "Mediation and the Paired-Associate Learning of Normals and Retardates." *American Journal of Mental Deficiency* 72: 610-13; January 1968.
- BROWN, LOUIS F., and RICE, JAMES A. "The Peabody Picture Vocabulary Test: Validity for EMRs." *American Journal of Mental Deficiency* 71: 901-903; May 1967.
- BUTTERFIELD, GAIL B., and BUTTERFIELD, EARL C. "The Effects of Training-Testing Interval and Intellectual Level Upon Transposition." *American Journal of Mental Deficiency* 72: 710-14; March 1968.
- CARTWRIGHT, G. PHILLIP. "Written Language Abilities of Educable Mentally Retarded and Normal Children." *American Journal of Mental Deficiency* 72: 499-505; January 1968.
- CAWLEY, JOHN F., and CHASE, DONNA V. "Productive Thinking in Retarded and Non-Retarded Children." *British Journal of Educational Psychology* 37: 356-60; November 1967.
- CHIAFFIN, JERRY D. *Production Rate as a Variable in the Job Success or Failure of Educable Mentally Retarded Adolescents*. Doctor's thesis. Lawrence: University of Kansas, 1967. 115 pp. Abstract: *Dissertation Abstracts* 28: 2432A-33A; No. 7, 1968.
- CHENNAULT, MADELYN. "Improving the Social Acceptance of Unpopular Educable Mentally Retarded Pupils in Special Classes." *American Journal of Mental Deficiency* 72: 455-58; November 1967.

- CLARKE, A. D. B., and COOPER, G. M. "Age and Perceptual-Motor Transfer in Imbeciles: Task Complexity as a Variable." *British Journal of Psychology* 57: 113-19; May 1966.
- CLAUSEN, JOHANNES. *Ability Structure and Subgroups in Mental Retardation*. New York: Spartan Books, 1966. 208 pp.
- COLE, SPURCEON N., and OTHERS. "Validity of an Abbreviated WISC Scale for Retarded and Borderline Children." *Psychological Reports* 21: 571-72; October 1967.
- CORTER, HAROLD M., and MCKINNEY, JAMES D. "Flexibility Training with Educable Retarded and Bright Normal Children." *American Journal of Mental Deficiency* 72: 603-609; January 1968.
- CROSSON, JAMES E. *The Experimental Analysis of Vocational Behavior in Severely Retarded Males*. Doctor's thesis. Eugene: University of Oregon, 1966. 140 pp. Abstract: *Dissertation Abstracts* 27: 3304A-3305A; No. 10, 1967.
- CROSSON, JAMES E. "A Technique for Programming Sheltered Workshop Environments for the Training of Severely Retarded Workers." *American Journal of Mental Deficiency*, 1968. (In press) (a)
- CROSSON, JAMES E. "The Functional Analysis of Behavior: A Technology for Special Education Practices." *Mental Retardation*, 1968. (In press) (b)
- DAVIS, LEO J., JR. "The Internal Consistency of the WISC with the Mentally Retarded." *American Journal of Mental Deficiency* 70: 714-16; March 1966.
- DE JUNG, JOHN E. "Modification of the Syracuse Scales of Social Relations for Administration in the Third and Fourth Grades and in Special Classes for the Mentally Retarded." Monograph Supplement 2-V18. *Psychological Reports* 18: 295-327; February 1966.
- DOUBROS, STEVE G., and DANIELS, GARY J. "An Experimental Approach to the Reduction of Overactive Behavior." *Behavior Research and Therapy* 4: 251-58; November 1966.
- EDMONSON, BARBARA, and OTHERS. "Increasing Social Cue Interpretations (Visual Decoding) by Retarded Adolescents Through Training." *American Journal of Mental Deficiency* 71: 1017-24; May 1967.
- ELKIN, LORNE. "Predicting Productivity of Trainable Retardates on Experimental Workshop Tasks." *American Journal of Mental Deficiency* 71: 576-80; January 1967.
- ELKIN, LORNE. "Predicting Performance of the Mentally Retarded on Sheltered Workshop and Non-Institutional Jobs." *American Journal of Mental Deficiency* 72: 533-39; January 1968.
- ELLIS, NORMAN R. "The Stimulus Trace and Behavioral Inadequacy." *Handbook of Mental Deficiency*. (Edited by Norman R. Ellis.) New York: McGraw-Hill Book Co., 1963. pp. 134-58.
- ELLIS, NORMAN R., editor. *International Review of Research in Mental Retardation*. New York: Academic Press, 1966. Vol. 1, 308 pp.; Vol. 2, 366 pp.
- ELLIS, NORMAN R., editor. *International Review of Research in Mental Retardation*. New York: Academic Press, 1968. Vol. 3, 288 pp.
- ELLIS, NORMAN R., and ANDERS, TERRY R. "Short-Term Memory in the Mental Retardate." *American Journal of Mental Deficiency* 72: 931-36; May 1968.
- FAGAN, JOSEPH F., III. "Short-Term Memory Processes in Normal and Retarded Children." *Journal of Experimental Child Psychology* 6: 279-96; June 1968.
- FANGMAN, THOMAS J. *Phrase Reading for the Mentally Retarded*. Doctor's thesis. Iowa City: University of Iowa, 1967. 98 pp. Abstract: *Dissertation Abstracts* 28: 3500A; No. 9, 1968.
- GARDNER, WILLIAM I. "Personality Characteristics of the Mentally Retarded: Review and Critique." *Behavioral Research in Mental Retardation*. (Edited by Herbert J. Prehm, Leo A. Hamerlynck, and James E. Crosson.) Eugene, Oreg.: Rehabilitation Research and Training Center in Mental Retardation, 1968. Monograph 1, pp. 53-68.
- GARDNER, WILLIAM I. "Use of the California Test of Personality with the Mentally Retarded." *Mental Retardation* 5: 12-16; August 1967.
- GARDNER, WILLIAM I., and BRANDL, CHARLENE. "Reinforcement Conditions and Incidental Learning in Mentally Retarded Adolescents." *American Journal of Mental Deficiency* 72: 215-19; September 1967.
- GATES, JOHN J., and FIXSEN, DEAN L. "Response Variability in Retarded Children." *Journal of Experimental Child Psychology* 6: 306-20; June 1968.



- GILES, DAVID K., and WOLF, MONTROSE M. "Toilet Training Institutionalized, Severe Retardates: An Application of Operant Behavior Modification Techniques." *American Journal of Mental Deficiency* 70: 766-80; March 1966.
- GOZALI, JOAV, and BIALER, IRV. "Children's Locus of Control Scale: Independence from Response Set Bias Among Retardates." *American Journal of Mental Deficiency* 72: 622-25; January 1968.
- GUSKIN, SAMUEL L., and SPICKER, HOWARD H. "Educational Research in Mental Retardation." *International Review of Research in Mental Retardation*. Vol. 3. (Edited by Norman R. Ellis.) New York: Academic Press, 1968. pp. 217-78.
- HAMERLYNCK, LEO A.; MARTIN, JACK W.; and ROLLAND, JOHN C. "Systematic Observation of Behavior: A Primary Teacher Skill." *Education and Training of the Mentally Retarded* 3: 39-42; February 1968.
- HAMMILL, DONALD D., and IRWIN, ORVIS C. "Factors Affecting Equivalency of PPVT and RSB When Used with Mentally Subnormal Children." *American Journal of Mental Deficiency* 71: 793-96; March 1967.
- HARTER, SUSAN. "Mental Age, IQ, and Motivational Factors in the Discrimination Learning Set Performance of Normal and Retarded Children." *Journal of Experimental Child Psychology* 5: 123-41; June 1967.
- HAWKINS, WILLIAM F. "Trace as a Predictor of Short-Term Memory of Organic and Familial Retardates." *American Journal of Mental Deficiency* 70: 576-79; January 1966.
- HEAL, LAIRD W. "The Role of Cue Value, Cue Novelty, and Overtraining in the Discrimination Shift Performance of Retardates and Normal Children of Comparable Discrimination Ability." *Journal of Experimental Child Psychology* 4: 126-42; October 1966.
- HEAL, LAIRD W.; ROSS, LEONARD E.; and SANDERS, BARBARA. "Reversal and Partial Reversal in Mental Defectives and Normal Children of a Comparable Mental Age." *American Journal of Mental Deficiency* 71: 411-16; November 1966.
- HEWETT, FRANK M.; MAYHEW, DONALD; and RABB, ETHEL. "An Experimental Reading Program for Neurologically Impaired, Mentally Retarded, and Severely Emotionally Disturbed Children." *American Journal of Orthopsychiatry* 37: 35-48; January 1967.
- HIMELSTEIN, PHILIP. "Use of the Stanford-Binet, Form L-M with Retardates: A Review of Recent Research." *American Journal of Mental Deficiency* 72: 691-99; March 1968.
- HUDDLE, DONALD D. "Work Performance of Trainable Adults Influenced by Competition, Cooperation, and Monetary Reward." *American Journal of Mental Deficiency* 72: 198-211; September 1967.
- JOHNSON, GORDON FLOYD. *An Investigation of Programed Procedures in Teaching Addition and Subtraction to Educable Mentally Retarded Subjects*. Doctor's thesis. Eugene: University of Oregon, 1966. 328 pp. Abstract: *Dissertation Abstracts* 27: 4132A; No. 12, 1967.
- JOHNSON, RICHARD F.; HAUGHTON, E.; and LAFAYE, HUGH G. "Behavior Therapy—Use in a Sheltered Workshop." *Diseases of the Nervous System* 26: 350-54; June 1965.
- KAREN, ROBERT L., and MAXWELL, SANDRA J. "Strengthening Self-Help Behavior in the Retardate." *American Journal of Mental Deficiency* 71: 546-50; January 1967.
- KAUFMAN, MELVIN E., and PREHM, HERBERT J. "A Review of Research on Learning Sets and Transfer of Training in Mental Defectives." *International Review of Research in Mental Retardation*. Vol. 2. (Edited by Norman R. Ellis.) New York: Academic Press, 1966. pp. 123-49.
- KELLAS, GEORGE, and BAUMEISTER, ALFRED A. "The Effects of Warning Signal Duration on the Reaction Times of Mental Defectives." *American Journal of Mental Deficiency* 72: 668-73; March 1968.
- KNAUS, WILLIAM J. *An Experimental Study of Three Methods of Programed Vocational Instruction Presented to Mentally Retarded Adolescent Females*. Doctor's thesis. Knoxville: University of Tennessee, 1967. 105 pp. Abstract: *Dissertation Abstracts* 28: 965A; No. 3, 1967.
- KOH, TONG-HE, and MADOW, ARNOLD A. "Relationship Between PPVT and Stanford-Binet Performance in Institutionalized Retardates." *American Journal of Mental Deficiency* 72: 108-13; July 1967.

- LELAND, HENRY, and OTHERS. "Adaptive Behavior: New Dimensions in the Classification of the Mentally Retarded." *Mental Retardation Abstracts* 4: 359-87; November 1967.
- LENT, JAMES R. "Mimosa Cottage: Experiment in Hope." *Psychology Today* 2: 51-58; June 1968.
- LEVINE, SAMUEL; ELZEY, FREEMAN F.; and PAULSON, FLOYD L. "Social Competence of School and Non-School Trainable Mentally Retarded." *American Journal of Mental Deficiency* 71: 112-15; July 1966.
- LILLIE, DAVID L. "The Effects of Motor Development Lessons on Mentally Retarded Children." *American Journal of Mental Deficiency* 72: 803-808; May 1968.
- LOVELL, K., and BRADBURY, B. "The Learning of English Morphology in Educationally Subnormal Special School Children." *American Journal of Mental Deficiency* 71: 609-15; January 1967.
- MCMANIS, DONALD L. "The von Restorff Effect in Serial Learning by Normal and Retarded Subjects." *American Journal of Mental Deficiency* 70: 569-75; January 1966.
- MCMANIS, DONALD L. "Marble-Sorting Persistence in Mixed Verbal-Incentive and Performance-Level Pairings." *American Journal of Mental Deficiency* 71: 811-17; March 1967.
- MAYER, C. LAMAR. "The Relationship of Early Special Class Placement and the Self-Concepts of Mentally Handicapped Children." *Exceptional Children* 33: 77-81; October 1966.
- MEYEROWITZ, JOSEPH H. "Peer Groups and Special Classes." *Mental Retardation* 5: 23-26; October 1967.
- MEYERS, ELIZABETH S., and MEYERS, C. E. "Problems in School Placement Arising from Divergent Conceptions of Educable Children." *Mental Retardation* 5: 19-22; October 1967.
- MILGRAM, NORMAN A. "Effects of Instructional Set on Verbal Conceptual Performance in Normal and Retarded Children." *Psychological Reports* 18: 647-53; April 1966. (a)
- MILGRAM, NORMAN A. "Verbalization and Conceptual Classification in Trainable Mentally Retarded Children." *American Journal of Mental Deficiency* 70: 763-65; March 1966. (b)
- MILGRAM, NORMAN A. "Retention of Mediation Set in Paired-Associate Learning of Normal Children and Retardates." *Journal of Experimental Child Psychology* 5: 341-49; September 1967.
- MILGRAM, NORMAN A., and FURTH, HANS G. "Factors Affecting Conceptual Control in Normal and Retarded Children." *Child Development* 38: 531-43; June 1967.
- MINGE, M. RONALD, and BALL, THOMAS S. "Teaching of Self-Help Skills to Profoundly Retarded Patients." *American Journal of Mental Deficiency* 71: 864-68; March 1967.
- NIHIRA, KAZUO; FOSTER, RAY; and SPENCER, LINDA. "Measurement of Adaptive Behavior: A Descriptive System for Mental Retardates." *American Journal of Orthopsychiatry* 38: 622-34; July 1968.
- OHLEICH, ELIZABETH S., and ROSS, LEONARD E. "Reversal and Nonreversal Shift Learning in Retardates as a Function of Overtraining." *Journal of Experimental Psychology* 72: 622-24; October 1966.
- PENNEY, RONALD K.; SEIM, ROBERT; and PETERS, RAY DEV. "The Mediational Deficiency of Mentally Retarded Children: I. The Establishment of Retardates' Mediational Deficiency." *American Journal of Mental Deficiency* 72: 626-30; January 1968.
- PETERSON, LINDA W. "Operant Approach to Observation and Recording." *Nursing Outlook* 15: 28-32; March 1967.
- PREHM, HERBERT J. "Associative Learning in Retarded and Normal Children as a Function of Task Difficulty and Meaningfulness." *American Journal of Mental Deficiency* 70: 860-65; May 1966. (a)
- PREHM, HERBERT J. "Retention in Retarded and Normal Children as a Function of Task Difficulty, Meaningfulness, and Degree of Original Learning." *Journal of Mental Deficiency Research* 10: 221-29; September 1966. (b)
- PREHM, HERBERT J. "Studies in Paired-Associate Learning: I. An Examination of Methodology." *American Journal of Mental Deficiency* 72: 492-95; November 1967.
- RICE, JAMES A., and BROWN, LOUIS F. "Validity of the Peabody Picture Vocabulary Test in a Sample of Low IQ Children." *American Journal of Mental Deficiency* 71: 602-603; January 1967.

- SCHOELKOPF, ANN M., and ORLANDO, ROBERT. "Reinforcement Delay Gradients of Retardates with a Concurrent Discrimination Task Procedure." *Psychological Record* 16: 113-28; April 1966.
- SEMMEI, MELVYN I., and WILLIAMS, JON. "Intentional and Incidental Learning in Normal, Borderline, and Retarded Children." *American Educational Research Journal* 5: 233-38; March 1968.
- SEMMEI, MELVYN I., and OTHERS. "A Grammatical Analysis of Word Associations of Educable Mentally Retarded and Normal Children." *American Journal of Mental Deficiency* 72: 567-76; January 1968.
- SEN, A. K.; CLARKE, ANN M.; and COOPER, G. M. "The Effect of Isolating Items in Serial Learning in Severely Retarded Subjects." *American Journal of Mental Deficiency* 72: 851-56; May 1968.
- SHAW, DALE J.; MATTHEWS, CHARLES G.; and KLOVE, HALLGRIM. "The Equivalence of WISC and PPVT IQs." *American Journal of Mental Deficiency* 70: 601-604; January 1966.
- SHEPARD, GEORGE. "Selected Factors in the Reading Ability of Educable Mentally Retarded Boys." *American Journal of Mental Deficiency* 71: 563-70; January 1967.
- SIEGEL, PAUL S., and FORMAN, GEORGE E. "Role of Incentive Strength in the Discrimination Learning of the Retardate." *Journal of Comparative and Physiological Psychology* 63: 552-54; June 1967.
- SILVERSTEIN, A. B. "Evaluation of a Split-Half Short Form of the WAIS." *American Journal of Mental Deficiency* 72: 839-40; May 1968.
- SKEELS, HAROLD M. *Adult Status of Children with Contrasting Early Life Experiences: A Follow-Up Study*. Monographs of the Society for Research in Child Development, Vol. 31, No. 3 (Serial No. 105). Chicago: University of Chicago Press, 1966. 65 pp.
- SMITH, ROBERT M. "Creative Thinking Abilities of Educable Mentally Handicapped Children in the Regular Grades." *American Journal of Mental Deficiency* 71: 571-75; January 1967.
- SPITZ, HERMAN H. "A Comparison of Mental Retardates and Normals on the Distorted Room Illusion." *American Journal of Mental Deficiency* 72: 34-39; July 1967.
- SPRADLIN, JOSEPH E.; GIRARDEAU, FREDERIC L.; and HOM, GEORGE L. "Stimulus Properties of Reinforcement During Extinction of Free Operant Response." *Journal of Experimental Child Psychology* 4: 369-80; December 1966.
- STEINMAN, WARREN M. "The Strengthening of Verbal Approval in Retardates by Discrimination Training." *Journal of Experimental Child Psychology* 6: 100-12; March 1968.
- STEPHENS, WYATT E. "Category Usage by Normal and Mentally Retarded Boys." *Child Development* 37: 355-61; June 1966. (a)
- STEPHENS, WYATT E. "Category Usage of Normal and Subnormal Children on Three Types of Categories." *American Journal of Mental Deficiency* 71: 266-73; September 1966. (b)
- SWITZER, MARY E. "The Coordination of Vocational Rehabilitation and Special Education Services for the Mentally Retarded." *Education and Training of the Mentally Retarded* 1: 155-60; December 1966.
- TERDAL, LEIF G. "Complexity and Position of Stimuli as Determinants of Looking Behavior in Retardates and Normals." *American Journal of Mental Deficiency* 72: 384-87; November 1967.
- VERGASON, GLENN A. "Recall of Paired-Associates in Retarded and Normal Subjects." *American Journal of Mental Deficiency* 71: 196-200; September 1966.
- WACHS, THEODORE D., and CROMWELL, RUE L. "Perceptual Distortions by Mentally Retarded and Normal Children in Response to Failure Information." *American Journal of Mental Deficiency* 70: 803-806; May 1966.
- WATSON, LUKE S., JR.; ORSER, RICHARD; and SANDERS, CHRISTOPHER. "Reinforcement Preferences of Severely Mentally Retarded Children in a Generalized Reinforcement Context." *American Journal of Mental Deficiency* 72: 748-56; March 1968.
- WEISS, DAVID, and WEINSTEIN, EUGENE. "Interpersonal Tactics Among Mental Retardates." *American Journal of Mental Deficiency* 72: 653-61; March 1968.
- WHITNEY, LINDA R. "Behavioral Approaches to the Nursing of the Mentally Retarded." *Nursing Clinics of North America* 1: 641-50; December 1966.
- WIESEN, ALLEN E., and OTHERS. "The Retarded Child as a Reinforcing Agent." *Journal of Experimental Child Psychology* 5: 109-13; March 1967.



- WINTERS, JOHN J., and GERJUOY, IRMA R. "Eye Movements and Verbal Reports in Tachistoscopic Recognition by Normals and Retardates." *Child Development* 38: 1193-99; December 1967.
- ZICLER, EDWARD; BUTTERFIELD, EARL C.; and GOFF, GENEVIEVE. "A Measure of Pre-institutional Social Deprivation for Institutionalized Retardates." *American Journal of Mental Deficiency* 70: 873-85; May 1966.

## CHAPTER II

### The Gifted

EDWARD C. FRIERSON

The present review appears at a significant time in the study of giftedness. One hundred years ago Galton published the first significant research on the gifted. One year ago Guilford (1967) presented his distinguished analysis of intelligence, which signaled the demise of the IQ concept after a 60-year reign. According to Parnes and Brunelle (1967), the research literature on creativity and giftedness during the year prior to Guilford's books equaled the amount reported during the previous decade, and this literature in turn equaled in volume the research reported from Galton's time.

The reader wishing to review the entire century of research findings in the area of the gifted would find the 1869-1959 period presented by Stein and Heinze in *Creativity and the Individual* (1960). Gowan's *Annotated Bibliography on Creativity and Giftedness* (1965) has provided a guide to the 1959-64 period. The research review offered by Gallagher and Rogge (1966) analyzed the 1963-65 period and cautiously predicted a research trend which has been fulfilled in a remarkable way. Said these authors, "One current emphasis seems to be on exploring the nature of creative thinking and the possibility of expanding the concept of superior intellectual ability itself." (p. 52)

Since 1965, research related to the gifted has indeed shifted dramatically from a concern for the gifted child to a concern for the creative process. So marked has been the trend away from studies concerning gifted children in favor of studies of creativity and cognitive factors that publications of associations for gifted children have experienced dwindling subscriptions and memberships while the new *Journal of Creative Behavior* (The Creative Education Foundation) has received wide acceptance since its introduction in 1967 and has more subscribers than all other related publications combined. Noffsinger (1968) confirmed that the research on the gifted published in professional journals from 1961-66 had changed its emphasis from gifted child studies (61 percent) in 1961 to creativity studies (64 percent) in 1966.

### Social Influences and Giftedness

Along with the shift in research emphasis among behavioral scientists, the "gifted-watchers" have observed a new social phenomenon developing among the gifted themselves during the past few years. The gifted are seek-

ing each other actively, are organizing, and are working collectively in many areas of life. The growth of MENSA, an organization requiring an IQ in the top 2 percent, has been remarkable (Serebriakoff, 1966). Over 11,000 Americans have joined this organization and have undertaken, among other things, the development of a new university of, by, and for the gifted. Through the publication *Mensa Bulletin* one may discover much about the gifted as MENSA members write about themselves.

The New Left, a political movement whose leadership includes many disenchanted intellectuals, has exposed the wide gap that still exists in our country between the typical "man-on-the-street" and the atypical "man-of-the-ivory-tower." The term *pseudointellectual* has become a political euphemism expressing contempt for the man who presumably is acting like an intellectual. Thus, the waves of political dissent have revealed openly the age-old attitudes of distrust, fear, and animosity which persons of different intellectual abilities often harbor toward one another. However, the rejection of contemporary values by student dissenters is seen by some researchers (e.g., Flacks, 1967) as a result rather than as a revolt.

Technology and racism have influenced our approach to creativity and giftedness during the past three years. While computer-simulated intelligence and the science of bionics revealed exciting new knowledge about man's bioneurological nature, investigators continued to search for deeply rooted ethno-cultural differences in intelligence (Boyer and Walsh, 1968; Kuttner and Lorincz, 1968). Both the technological revolution and the sociological revolution seem to have obscured individual uniqueness in favor of dehumanized analyses on the one hand and superficial group identity on the other.

Another significant social influence during the past few years has been the impact of drugs on human experience. Efforts to expand one's consciousness have ranged from hippie-oriented escapades to controlled medical experimentation. The study of biochemical agents and creative behavior is in its infancy but the potential knowledge inherent in such study may revolutionize therapeutic procedures involving the potentially gifted (Hess, 1967; Krippner, 1968; Ludwig, 1966).

The influences discussed thus far suggest that researchers no longer are investigating educational matters relative to the gifted. However, this review introduces recent research findings which bear heavily upon the education of gifted children.

### The Creativity Research

In a graphic description, "Creativity—A Blind Man's Report on the Elephant," Yamamoto (1965) reported the confusion among creativity researchers. He demonstrated that philosophical differences among research workers were responsible for failure to agree about creativity. Studies in the past three years have continued to reflect (a) different points of depar-



ture in the definition of creativity, (b) different assumptions and presuppositions about creativity, and (c) different research strategies among researchers with different orientations. In addition, the criterion problem has influenced creativity research throughout the mid-1960's (White and Williams, 1965).

Researchers continued to investigate whether creativity is independent of the general intelligence factor. Clark, Veldman, and Thorpe (1965) reported a creativity-intelligence dichotomy consistent with the earlier work of Getzels and Jackson (1962), although Marsh (1964) had contradicted the Getzels-Jackson position with his analyses and had concluded that the conventional IQ remained the best single criterion for creative potential. The confusion remained as Cicirelli (1965) added that creativity did not interact with IQ to boost achievement, while the research by Schmadel, Merrifield, and Bonsall (1965) suggested that children of high IQ may be both gifted and creative.

The research of Wallach and Kogan (1965) highlighted the period as they differentiated modes of thinking in children and introduced a single dimension or characteristic of intelligence which could be considered a facet of creativity. These investigators defined individual difference in the ability to produce many unique cognitive associates as this characteristic. Ten measures of this ability proved to be highly intercorrelated while correlations between these measures and intelligence tests proved to be very low. Subsequently, the new dimension—ability to generate unique and plentiful associates—was used by Wallach and Kogan to investigate group differences in personality, self-concept, fantasy, and conceptual abilities among young children.

Wallach and Kogan specified the term *creativity* by defining it as the ability to generate unique and plentiful associates. However, this single ability may prove more popular with researchers than with educators. The multifactor creativity test battery developed by Torrance (1966) has been widely used for research and holds promise for changing classroom instruction on behalf of the creatively gifted.

*Additional References:* Gowan, Demos, and Torrance (1967); Joncich (1964); Kneller (1966); Mooney and Razik (1967); Yamamoto (1965a).

## Identifying the Gifted

Perhaps the most intriguing measurement study in the past few years was reported by John Ertl, director of the Center of Cybernetic Studies at the University of Ottawa (Ertl, 1968). He presented findings showing a high degree of relationship between the electrical activity of the human brain and intelligence as measured by psychological tests. One hundred experimental subjects, ages 18-50, were given one of the Otis tests of intelligence. The IQ scores ranged from 77 to 136. Evoked potentials (changes

in electrical activity of the brain in response to sensory stimulation, such as flashing lights) were detected by EEG, were averaged by a computer, and were photographed (polaroid pictures of oscillograms). The experimenter found that *the higher the intelligence, the shorter the response time*. Since it is very difficult to produce a change in the evoked pattern of an individual, a stable new measure of intellectual potential may be forthcoming. The educational significance of identifying a valid hereditary factor, Ertl's neurological efficiency, has not been determined.

Martinson and Seagoe (1967) attempted to differentiate the gifted by evaluating the creative products of schoolchildren. The researchers compared the scores from the *Stanford-Binet Intelligence Scale* and the Guilford tests of divergent thinking with judges' ratings of creative products in the areas of science, writing, social studies, music, and art. Interjudge reliability for the various products was extremely mixed, ranging from 0.95 in judging clay products to 0.14 in judging poetry. Comparisons between high and low IQ groups demonstrated a significant correlation between judges' ratings and *Stanford-Binet* scores. The high IQ groups (mean: 142.7) produced more creative products than did the low IQ group (mean: 107.5). Since the *Stanford-Binet* scores were unrelated to divergent thinking test scores, a familiar question was raised, "What do divergent thinking tests measure?" The traditional IQ proved to be a better predictor of creative productions!

A new originality measure, the Preconscious Activity Scale, was developed by Holland and Baird (1968) on the basis of personality study rather than product or task analysis. The scale was derived from Kubie's book *Neurotic Distortion of the Creative Process* (1958). Thirty-eight true-false items comprise the scale which was administered to more than 7,000 college men and women in validation studies. High scorers had artistic, intellectual, and feminine interests and accomplished notable achievements requiring persistence, high level skills, and originality (Holland and Richards, 1965; Richards, Holland, and Lutz, 1967).

The *Torrance Tests of Creative Thinking* have emerged as the most influential new measures in identifying the gifted. However, Torrance (1967), in summarizing 300 studies in which the tests were used, expressed concern that, while the studies continued to investigate the relationship between measures of creative behavior and measures of intelligence, they have "been too superficial to add much to our understanding" of the dynamics of this relationship.

*Additional References:* Khatena and Gowan (1967); Mednick and Andrews (1967); Yamamoto (1965b).

## Teaching the Gifted

New school arrangements such as ungraded continuous progress programs, supplementary learning centers, team teaching, and mobile teaching

laboratories have provided opportunities for education to employ new instructional approaches with students. Unfortunately, little innovation in curriculum has been based on knowledge of giftedness. The research presented in this section explored new teaching materials, procedures, and objectives for the gifted.

In an attempt to stimulate creative thinking abilities, Parnes (1966) introduced a programed, creative problem-solving course. Along with changes in creative ability, the researcher studied the attitudes of students toward programed and instructor-led courses. Three groups of 62 college-bound students were matched on IQ and were identified as control group, program group, and instructor-led group. Six schools were involved in order to eliminate discussion among groups. Groups met during study periods and were required to perform no outside work. Eleven pretests were administered to each group after which a 13-week course, with two sessions each week, was offered. Both groups of experimental subjects made greater gains on creativity measures than did the control subjects. Those students who were instructor-led were superior to students who used the program alone. Instructor-led students found the course more interesting, although both experimental groups reported equal application of what they had done. Findings suggested that teachers can teach students to approach problems creatively and that gifted students can develop new problem-solving skill through autoinstructional materials.

Hauck (1967) reported improvement in evaluative thinking abilities among gifted sixth-grade students who received special instruction for one month. Unique lesson plans were developed using three psychological constructs: mental set, game-model theory, and feedback. Lessons were 30 minutes each and were designed to have application in all curricular areas. Students assumed an evaluator role each day (secretary evaluating a speech for employer, policeman evaluating aspects of a situation involving crime, etc.), completed seatwork tasks, and participated in specified group activities. Teacher feedback was directed toward thinking processes rather than toward the responses or products of thinking. Using several Guilford tests of evaluative thinking, Hauck indicated that evaluative abilities were improved significantly during the month of lessons and the gains were increased over a 30-day follow-up period. Since no special instruction was introduced during the follow-up period, the findings suggest that once understood, styles of thinking and problem-solving approaches were employed by gifted children without special cues. In spite of the limitations arising from the use of measures not standardized for young children, the application of Guilford's (1967) structure of intellect concept to regular classroom practices was supported by this study.

Beittel (1966) reported an unusual and successful use of instructional media with art students. Students were involved in a four-week experiment which required a 75-minute period each week. While students worked, their drawings were photographed every three minutes. Completed works were



photographed at the end of each period. Through the use of the photographs, delayed evaluation of products was accomplished in an environment where the self-reflection phase of learning could be separated from the activity phase. Theoretically, the self-reflective phase in art activity is the only one in which learning in art can be influenced without violating the privacy of the art dialogue. The experimental group exhibited gains in art through the delayed self-reflection treatment. Beittel admitted the shortcomings of techniques such as the stop-action photographs and commented that "more theory and less gadgetry is the urgent need."

The use of television shows and movies especially developed for the gifted increased during the past few years. Notable among research with these media was the work of Drews (1965). Working with 243 talented ninth-graders, Drews produced 10 half-hour films, *The Being and Becoming Series*, which dealt with the life-styles and value systems of creative and socially concerned adult prototypes. The films were augmented by a flexible anthology which introduced students to significant ideas and crucial issues.

The experimental group met 50 minutes a day for a semester while a control group was given a traditional career-guidance course. All subjects were administered a personality inventory. The Allport-Vernon-Lindzey *Study of Values*, a Critical Thinking Test, the Rokeach Dogmatism Scale and Rigidity Scale, along with informal student-interest surveys, an occupational-choice survey, a profile check, and an Ability Self-Concept Rating. Eleven post-tests were administered. The experimental group scored higher on the Originality and Theoretical Orientation Scales of the personality inventory and higher on the Estheticism and Complexity Scales. Experimental subjects also scored higher on the Theoretical Scale of the *Study of Values* and the ACE Critical Thinking Test. Experimental subjects checked more creative intellectual reasons for choice of future occupation. In short, attitudes changed in the direction of greater motivation to learn, more openness to psychological growth, and a greater orientation toward social concern. This research demonstrated that personality and attitude changes can be effected among the gifted when teaching procedures emphasize goals other than the traditional achievement objectives.

*Additional References:* Berkowitz (1967); Hutchinson (1967); Neff (1967).

### Developing Creativity

Three factors which affect the development of creative and productive thinking were studied recently by researchers. Bishop (1968) analyzed successful teachers of the gifted. Gallagher, Aschner, and Jenne (1967) scored teacher-student interactions and related productive thinking responses to family, personality, attitude, and cognitive traits. Torrance (1965) varied evaluation techniques and observed the differential effect which certain reinforcement had on student originality.

One hundred and nine teachers who most influenced gifted secondary students prior to their participation in a statewide honors program were compared with 97 teachers who were not identified as influential (Bishop, 1968). Influential teachers did not differ from unselected teachers in sex, marital status, type of undergraduate institution, highest degree held, course work preparation, or professional activity. The selected teachers were intellectually superior, ranking in the top 3 percent on adult population norms. These teachers pursued intellectual avocations, were more active in the cultural life of their communities, and exhibited a higher need to achieve. The successful teachers exhibited more favorable attitudes toward students and took a greater interest in student motives, feelings, and behaviors. These teachers were more considerate of student opinions in class; were most systematic, orderly, and businesslike in their approach; and exhibited more imagination and enthusiasm about their particular subject. They also supported special education provisions for gifted students and more preferred to teach a class of exceptionally bright students than did their fellow teachers.

In the Gallagher, Aschner, and Jenne studies (1967), the Aschner-Gallagher Classification System was used to assess different types of cognitive behavior expressed in the classroom. Cognitive-memory questions made up more than 50 percent of the questions asked by teachers. In some sessions, requests for divergent and evaluative thinking were entirely absent. Significant differences were noted among teachers both in cognitive style and in the type of questions asked. Teachers' performances varied significantly from class section to section even when identical concepts were taught. In the final analysis, the evidence indicated that intellectual productivity in school is directly related to the teacher's style, expectation, and response pattern.

The effects of teacher evaluation on creative thinking was contrasted with the effects of peer evaluation by Torrance (1965). In one study, teachers offered a two-dollar prize for stories which, under condition A, had correct spelling, punctuation, grammar, neatness, margins, indentations, and handwriting; and under condition B, were the most interesting, exciting, unusual, and original. In another study, children were placed in an experimental condition following practice periods during which some received no evaluation and some received appropriate direction and suggestions. Critical and creative peer-evaluation conditions were developed in a third study by encouraging students to point out things that went wrong with products and to suggest possibilities for making products more unusual, interesting, complete, or elaborate.

Children achieved best along those lines for which they were rewarded. When rewarded for correctness, children produced few errors. When rewarded for original and interesting work, children produced significantly higher ratings on those qualities. Young children under nonevaluative practice conditions exhibited a higher degree of creative production, but

among older children no significant differences were found. Creative peer evaluation produced especially noticeable differences in creativity among older children. Along with teacher characteristics and teacher style, teacher evaluation seemed to contribute directly toward differences in creative and productive thinking gains in children.

*Additional References:* Gowan, Demos, and Torrance (1967); Hutchinson (1967); Walker (1966).

## Educational Practices

Regardless of program innovations, educational procedures for the gifted continued to be influenced by traditional grading practices and standardized testing. Hausdorff and Farr (1965) examined the effect of grading practices on the grades received by gifted students. The grades from school districts having a single grading system (grading based upon heterogeneous group achievement standard) were compared with grades from districts having a dual system (grading based upon heterogeneous group achievement standard and upon homogeneous group standard). Gifted students received significantly lower grades on report cards when graded on homogeneous standards. However, there were no differences among students on standardized achievement tests. This descriptive research supported the complaint that the gifted child who is marked on homogeneous group standards fares poorly in terms of the grades which become part of his permanent record. Using a chi-square analysis, the authors found significant differences in grades between groups in all subjects except spelling.

Howell (1965) matched 28 high school honors students (mean IQ: 142) with 21 nonhonors students (mean IQ: 135). As seniors, all students took the College Board examinations. The honor students scored significantly higher on the verbal test. The superiority of the honors group was attributed to the grouping and the enriched curriculum which these students enjoyed. It was further claimed that the grouping increased the competition, which in turn created a motivational climate not found in the control group. No evidence for these claims was presented.

Both studies continued to demonstrate the weakness of justifying provisions for the gifted in traditional terms. Student change in special enrichment programs at the elementary level and secondary level reported in terms of grades and College Board examination scores had limited meaning.

A promising effort to describe more comprehensively the impact of curricular adaptations was described by Meeker (1968). She included measures of creativity, leadership, and motor skills as well as grades, intelligence ratings, and achievement scores in a four-year follow-up study of 67 gifted students. These students were enrolled in different high school programs (one experimental, the other traditional) after having participated in a specially designed elementary school program for the gifted. Meeker found



that potentially brilliant achievers were lost as a consequence of inadequate attention during the secondary school years. She attributed poor records to stereotyped expectations and dedication to traditional departmentalized practices.

*Additional References:* Parnes (1968); Taylor (1967); Torrance and Witt (1966).

## Differences Among the Gifted

Differences in creativity among children were linked in recent studies to racial (caste) differences, socioeconomic (class) differences and general-intelligence (IQ) differences. Smith (1965) explored the relationship of creativity to social class in a study which involved 359 Negro and 244 white children. He found that among white subjects children from higher socioeconomic circumstances scored above children from lower socioeconomic levels on several verbal creativity measures. With IQ covaried, the lower socioeconomic subjects scored better, though not significantly, than the higher socioeconomic subjects on nonverbal creativity measures. There were no significant differences between middle and lower class Negro subjects.

Gifted subjects (IQ: 120+) in the Smith study exceeded normal subjects on 8 of the 14 verbal creativity variables at an 0.05 level of significance. None of the nonverbal factors yielded a significant group difference. Differences in creativity between Negro subjects and white subjects were significant although disproportionate socioeconomic class samples exaggerated differences. It was suggested that serious attention must be given to the possibility that a caste system, no matter how subtle, might exert a stronger influence on style of life or style of thinking than a class system.

Differences between gifted children from upper and lower status backgrounds were also examined by Frierson (1965). Measures of interests, activities, personality traits, creativity, stature, and school achievement were administered to gifted children enrolled in the Cleveland Major Work Program. Upper status gifted children differed from lower status gifted children in the quantity and quality of their reading, in their positive attitudes toward school, in their lower preference for competitive team sports, and in their awareness of parental aspirations for college attendance. Verbal measures of creativity favored the upper socioeconomic group. The gifted groups were found to be more similar in personality patterns, height, weight, and interest responses than were groups of gifted and nongifted children from the same socioeconomic background.

Barbe (1964) studied the personal adjustment, family background, and school achievement of 65 highly gifted children (mean IQ: 158) and 65 moderately gifted children (mean IQ: 129). The sample, drawn from throughout the State of Ohio, included 31 matched pairs of boys and 34

matched pairs of girls. The highly gifted subjects were found in the more affluent and the more highly educated families. The highly gifted group scored significantly higher on four out of five creativity tasks. No differences were reported in beginning age for learning to read or in reaching physical milestones. The highly gifted group was reported as having more behavior irregularities. Only one personality trait, surgency, was significantly different between groups. The moderately gifted group tended toward a more serious and less mercurial outlook.

Differences between achieving and underachieving gifted secondary level boys were outlined by Raph, Goldberg, and Passow (1966). Students in two settings, Evanston (Illinois) Township High School and Dewitt High School in New York, were identified as underachieving if they had an IQ score above 125 and had a school grade average which placed them at or below the average for the general student population. Two groups of bright underachievers in the Dewitt High School setting were placed in a special group guidance and study skills class. A third group acted as a control. No differences in grade averages were found among the three groups at the end of grade 10, grade 11, or grade 12. The groups did not differ on number of discharges, dropouts, failures, repeaters, or scholarship winners. By the end of grade 10, the gap between the underachievers and the achievers was greater than it had been at the beginning of the year or at midyear. The intervention showed little promise when initiated at the high school level.

## Overview

The period 1965-68 saw a substantial shift in research emphasis from the nature of gifted children to the nature of creativity. Societal influences affecting research on the gifted in this period ranged from the New Left political movement to computer technology, racism, and consciousness-altering drugs. Experimental teaching approaches were introduced during the period and curricular modifications were evaluated. New identification techniques were suggested and differences among the gifted were reported. The period reflected that the sheer bulk of creativity-related research each year now exceeds that quantity of research on the gifted conducted during the entire 1850-1950 period.

## Bibliography

- BARBE, WALTER B. *One in a Thousand: A Comparative Study of Highly and Moderately Gifted Elementary School Children*. Columbus: Ohio State Board of Education, 1964, 78 pp.
- BEITTEL, KENNETH R. "Instructional Media for Creativity in the Arts." *Instructional Media and Creativity*. (Edited by Calvin W. Taylor and Frank E. Williams.) New York: John Wiley and Sons, 1966. pp. 277-307.

- BISHOP, WILLIAM E. "Successful Teachers of the Gifted." *Exceptional Children* 34: 317-25; January 1968.
- BOYER, WILLIAM H., and WALSH, PAUL. "Are Children Born Unequal?" *Saturday Review* 51: 61-63, 77-79; October 1968.
- CICIRELLI, VICTOR G. "Form of the Relationship Between Creativity, IQ, and Academic Achievement." *Journal of Educational Psychology* 56: 303-308; December 1965.
- CLARK, CHARLES M.; VELDMAN, DONALD J.; and THORPE, JOSEPH S. "Convergent and Divergent Thinking Abilities of Talented Adolescents." *Journal of Educational Psychology* 56: 157-63; June 1965.
- DREWS, ELIZABETH M. *The Effectiveness of Special Training with Audiovisuals in Changing Aspirations of Intellectually Superior Students*. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 6471. East Lansing: Michigan State University, 1965. 285 pp.
- ERTL, J. P. "Intelligence Testing by Brainwaves." *Mensa Bulletin* 110: 3-6; April 1968.
- FLACKS, RICHARD. "Student Activists: Result, Not Revolt." *Psychology Today* 1: 18-23, 61; October 1967.
- FRIERSON, EDWARD C. "Upper and Lower Status Gifted Children: A Study of Differences." *Exceptional Children* 32: 83-90; October 1965.
- GALLAGHER, JAMES, and ROGGE, WILLIAM. "The Gifted." *Review of Educational Research* 36: 37-55; February 1966.
- GALLAGHER, JAMES; ASCHNER, MARY JANE; and JENNE, WILLIAM. *Productive Thinking of Gifted Children in Classroom Interaction*. Washington, D.C.: Council for Exceptional Children, 1967. 103 pp.
- GETZELS, JACOB W., and JACKSON, PHILLIP W. *Creativity and Intelligence*. New York: John Wiley and Sons, 1962. 293 pp.
- GOWAN, JOHN C. *Annotated Bibliography on Creativity and Giftedness*. Northridge, Calif.: San Fernando Valley State College Foundation, 1965. 197 pp.
- GUILFORD, J. P. *The Nature of Human Intelligence*. New York: McGraw-Hill Book Co., 1967. 538 pp.
- HAUSDORFF, HENRY, and FARR, S. DAVID. "The Effect of Grading Practices on the Marks of Gifted Sixth Grade Children." *Journal of Educational Research* 59: 169-72; December 1965.
- HAUCK, BARBARA B. "A Comparison of Gains in Evaluation Ability Between Gifted and Nongifted Sixth Grade Students." *Gifted Child Quarterly* 11: 166-71; Autumn 1967.
- HESS, WALTER R. "Causality, Consciousness, and Cerebral Organization." *Science* 158: 1279-83; December 1967.
- HOLLAND, JOHN L., and BAIRD, LEONARD L. "The Preconscious Activity Scale: The Development and Validation of an Originality Measure." *Journal of Creative Behavior* 2: 217-25; Summer 1968.
- HOLLAND, JOHN L., and RICHARDS, JAMES M., JR. "Academic and Nonacademic Accomplishment: Correlated or Uncorrelated?" *Journal of Educational Psychology* 56: 165-74; August 1965.
- HOWELL, WALLACE J. "Influence of Curriculum Enrichment in a High School Honors Group on College Board Examination Scores." *Journal of Educational Research* 59: 113-14; November 1965.
- KRIPPNER, STANLEY. "The Psychedelic State, the Hypnotic Trance, and the Creative Act." *Journal of Humanistic Psychology* 8: 49-67; Spring 1968.
- KUBIE, LAWRENCE S. *Neurotic Distortion of the Creative Process*. Lawrence: University of Kansas Press, 1958. 151 pp.
- KUTTNER, ROBERT E., and LORINCZ, ALBERT B. "Utilization of Accentuated Environmental Inequalities in Research on Racial Differences." *Science* 160: 439-40; April 1968.
- LUDWIG, ARNOLD M. "Altered States of Consciousness." *Archives of General Psychiatry* 15: 225-34; September 1966.
- MARSH, R. W. "A Statistical Re-Analysis of Getzels and Jackson's Data." *British Journal of Educational Psychology* 34: 91-93; February 1964.
- MARTINSON, RUTH A., and SEAGOE, MAY V. *The Abilities of Young Children*. Washington, D.C.: Council for Exceptional Children, 1967. 66 pp.
- MEEKER, MARY. "Differential Syndromes of Giftedness and Curriculum Planning: A Four Year Follow-Up." *Journal of Special Education* 2: 185-94; Winter 1968.



- NOFFSINGER, THOMAS. "Creativity: A Critique." *Science Education*, 1968. (Accepted for publication)
- PARNES, SIDNEY J. *Programming Creative Behavior*. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 5-0716. Buffalo: New York State University, 1966. 134 pp.
- PARNES, SIDNEY J., and BRUNELLE, EUGENE A. "The Literature of Creativity, Part I." *Journal of Creative Behavior* 1: 52-109; Winter 1967.
- RAPH, JANE B.; GOLDBERG, MIRIAM L.; and PASSOW, A. HARRY. *Bright Underachievers*. New York: Teachers College Press, 1966. 289 pp.
- RICHARDS, JAMES M., JR.; HOLLAND, JOHN L.; and LUTZ, SANDRA W. "Prediction of Student Accomplishment in College." *Journal of Educational Psychology* 58: 343-55; December 1967.
- SCHMADEL, ELNORA; MERRIFIELD, PHILLIP R.; and BONSALE, MARCELLA R. "A Comparison of Performances of Gifted and Nongifted Children on Selected Measures of Creativity." *California Journal of Educational Research* 16: 123-28; May 1965.
- SEREBRIAKOFF, VICTOR. *IQ: A Mensa Analysis and History*. London: Hutchinson and Co., 1966. 192 pp.
- SMITH, ROBERT M. *The Relationship of Creativity to Social Class*. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 2250. Pittsburgh: University of Pittsburgh, School of Education, 1965. 141 pp.
- STEIN, MORRIS J., and HEINZE, SHIRLEY J. *Creativity and the Individual*. Glencoe, Ill.: Free Press, 1960. 428 pp.
- TORRANCE, E. PAUL. *Rewarding Creative Behavior*. Englewood Cliffs, N.J.: Prentice-Hall Co., 1965. 353 pp.
- TORRANCE, E. PAUL. *Torrance Tests of Creative Thinking*. Princeton, N.J.: Personnel Press, 1966. Manual, 95 pp.
- TORRANCE, E. PAUL. "The Minnesota Studies of Creative Behavior: National and International Extensions." *Journal of Creative Behavior* 1: 137-54; Spring 1967.
- WALLACH, MICHAEL A., and KOGAN, NATHAN. *Cognitive Originality, Physiognomic Sensitivity, and Defensiveness in Children*. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 1316B. Durham, N.C.: Duke University, 1965. 217 pp.
- WHITE, WILLIAM F., and WILLIAMS, ROBERT E. "Identification of Creativity and the Criterion Problem." *Journal of Secondary Education* 40: 275-81; October 1965.
- YAMAMOTO, KAORU. "Creativity—A Blind Man's Report on the Elephant." *Journal of Counseling Psychology* 12: 428-34; Winter 1965.

### Additional References

- BERKOWITZ, HARRY. "An Alternative Position in Preparing Talented Tots: Some Problems and Implications." *Gifted Child Quarterly* 11: 241-50; Winter 1967.
- GOWAN, JOHN C.; DEMOS, GEORGE G.; and TORRANCE, E. PAUL. *Creativity: Its Educational Implications*. New York: John Wiley and Sons, 1967. 336 pp.
- HUTCHINSON, WILLIAM L. "Creative and Productive Thinking in the Classroom." *Journal of Creative Behavior* 1: 419-27; Fall 1967.
- JONCICH, GERALDINE. "A Culture-Bound Concept of Creativity." *Educational Theory* 14: 133-43; July 1964.
- KHATENA, JOE, and GOWAN, J. C. "Crosscultural Measurement of Intelligence with the DAM and CPM." *Gifted Child Quarterly* 11: 227-30; Winter 1967.
- KNELLER, GEORGE F. *The Art and Science of Creativity*. New York: Holt, Rinehart, and Winston, 1966. 106 pp.
- MEDNICK, MARTHA T., and ANDREWS, FRANK W. "Creative Thinking and Level of Intelligence." *Journal of Creative Behavior* 1: 428-31; Fall 1967.
- MOONEY, ROSS L., and RAZIK, TAHER A. *Explorations in Creativity*. New York: Harper and Row, 1967. 338 pp.
- NEFF, HERBERT B. "Ways To Help Gifted Students Like Social Studies." *Gifted Child Quarterly* 11: 108-11; Summer 1967.
- PARNES, SIDNEY J. "Methods and Educational Programs for Stimulating Creativity: A Representative List." *Journal of Creative Behavior* 2: 71-75; Winter 1968.

- TAYLOR, CALVIN W. "Questioning and Creating: A Model for Curriculum Reform." *Journal of Creative Behavior* 1: 22-33; Winter 1967.
- TORRANCE, E. PAUL, and WITT, GEORGE. "Experimental Grouping on the Basis of Creative Abilities and Motivations." *Gifted Child Quarterly* 10: 9-14; Spring 1966.
- WALKER, W. J. "Creativity and High School Climate." *Gifted Child Quarterly* 10: 139-44; Autumn 1966.
- YAMAMOTO, KAORU. *Creative Thinking Abilities and Peer Conformity in Fifth Grade Children*. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 2021. Kent, Ohio: Kent State University, 1965. 225 pp. (a)
- YAMAMOTO, KAORU. "Validation of Tests of Creative Thinking: A Review of Some Studies." *Exceptional Children* 31: 281-90; February 1965. (b)

## CHAPTER III

### The Speech Handicapped

CLARK STARR

This chapter contains reviews of selected research publications that appeared between January 1, 1965, and October 1, 1968. It is divided into sections on articulation disorders, stuttering, voice disorders, delayed speech and language, and aphasia.

#### Articulation Disorders

Articulation disorders continue to be the central concern of many clinical speech pathologists. At the present time, there are no adequate theories or models to account for the acquisition or maintenance of articulation. Researchers appear to derive their questions from clinical observations and needs. Extensive efforts have been made to identify motor, sensory, and perceptual differences between subjects (Ss) with normal and disordered articulation and to discover prognostic indicators for therapy.

#### Motor, Sensory, and Perceptual Factors

Bosma (1967) edited a volume containing a series of studies on oral sensation and perception. Several studies used geometric forms devised by Bosma to test oral stereognosis in Ss with articulation disorders. Findings were contradictory, but they suggest that stereognosis is not a sufficiently strong variable to account for most articulation disorders. In collaboration with a number of others, Ringel (1965, 1967, 1967) presented extensive data on two-point discrimination, mandibular kinaesthesia, and texture discrimination in the speech mechanism. The significance of these data for theoretical or clinical purposes has not been tested.

Shelton and others (1966) tried unsuccessfully to predict articulation proficiency from observations of motor behavior of the speech mechanism during nonspeech activities. Palmer, Wourth, and Kincheloe (1964) found lingual labial apraxia in Ss diagnosed as having functional articulation problems, and Jenkins and Lohr (1964) found that children with functional articulation problems scored lower than their controls on the *Oseretsky Tests of Motor Proficiency*.

In separate X-ray studies, Subtelny, Mestre, and Subtelny (1964) and Weinberg (1968) found that Ss with malocclusions who articulate /s/ adequately maintain the same relationships between tongue tip and lower



incisors as do Ss with normal occlusion and adequate articulation. Snidecor and Karies (1965) used dental appliances to correct open bites for Ss with defective sibilants and found articulation changed but not necessarily improved. Heighton and Nichols (1967) studied children with open bites and children with normal occlusions and found the former group to have less adequate articulation than the latter group.

Ronson (1965) reported "abnormal swallowing habits" to be associated with lisping. Fletcher and Meldrum (1968) found a relationship between length of the lingual frenulum and articulation proficiency.

Haller (1967) reported that Ss with functional articulation disorders performed more poorly than their controls on speech sound discrimination tasks administered under unfavorable listening conditions. Hutchinson and Clark (1968) examined four auditory discrimination tests and concluded that self-monitoring tests were the most sensitive in differentiating between Ss with normal articulation as opposed to those with defective articulation. Christine and Christine (1964) found speech sound discrimination scores related to reading as well as to articulation. Rechner and Wilson (1967) found low sound discrimination scores associated with low scores on the *Illinois Test of Psycholinguistic Abilities (ITPA)* subtests concerned with auditory memory span, tenses, and plurals. Sherman and Geith (1967) studied Ss with high and low speech sound discrimination scores and found the high group to have better articulation and higher scores on the *Peabody Picture Vocabulary Test (PPVT)*. Weiner (1967) reviewed available studies and concluded that a positive relationship between speech sound discrimination and articulation is most likely to occur in Ss who are under nine years of age and have severe articulation disorders.

### Description and Diagnosis of Articulation Disorders

Smith and Ainsworth (1967) reported that Ss responses to articulation tests using pictures as stimuli contained more errors than their responses to auditory or auditory-visual tests. The latter tests produced the fewest errors. Nichols (1967) evaluated three published and five unpublished articulation screening tests and found them all reliable, but varied in their pass-fail criteria. He concluded that time and expense were basic factors to be considered in choosing a screening test. Gilmore and Familant (1967) found some pictures in the widely used *Templin-Darley Screening and Diagnostic Tests of Articulation* to be inappropriate for eliciting responses from a lower socioeconomic population in the South.

Prins and Bloomer (1968) developed a consonant intelligibility test. They reported that listeners responded reliably and that the test differentiated a wide range of speakers with cleft palates and articulation defects. Korst (1966) found high correlations between the *PPVT* and the *Leiter International Performance Scale* and suggested that the *PPVT* was usable for screening Ss with severe speech problems.

## Therapy for Articulation Disorders

Wilson (1966) studied the effects of articulation therapy in a large group of educable mental retardates being treated in a public school setting and found therapy to be ineffective in modifying articulation. Stoia and Reeling (1967) reported that Ss participating in a Head Start Program improved their articulation test scores more than the control Ss did. Sommers and others (1966) presented data suggesting that, in the school setting, group therapy is as effective as individual therapy for children with mild and moderately severe articulation disorders.

Elbert, Shelton, and Arndt (1967) and Shelton, Elbert, and Arndt (1967) devised a procedure for assessing the acquisition of new articulation responses and evaluating therapy on a session-to-session basis. They found their procedure to be sensitive and reliable. Using their procedure they found that therapy directed toward the improvement of /s/ resulted in improvement of its voiced cognate /z/ but not in improvement of an unrelated phoneme /r/. In a subsequent study, they showed that Ss who improved their /r/ production during therapy maintained improvement on a stable basis after therapy was discontinued.

Baer and Winitz (1968) studied the relationship between inconsistency of misarticulation and speech sound learning. They found that, for the phoneme /v/, inconsistency did not seem to be highly correlated with speech sound learning. Sommers and others (1967) described a study in which therapy was found to be more effective for Ss who responded poorly to direct stimulation tests than for Ss who responded well. Winitz and Preisler (1967) found pretraining Ss on distinctive features of consonants aided them in learning difficult speech sound discriminations. Lichtenberg (1966) concluded that intensive auditory training improved speech sound discrimination test scores for hard of hearing Ss and Ss with normal hearing and good articulation but not for Ss with normal hearing and defective articulation.

Griffith (1965) reported successful modification of /s/ and /r/ using instrumental conditioning. An extensive discussion of research related to operant procedures and articulation was presented in a book edited by Sloane and MacAulay (1968).

## Stuttering

In recent literature, the application of operant conditioning procedures to the analysis of stuttering has been a central concern, along with linguistic analysis of stutterers' verbal output. However, traditional interest in the adaptation phenomenon has not diminished.

Martin (1965) reported that lay and professional listeners judge stuttering severity as being greater when presented with audiovisual stimuli than when presented with audio stimuli alone. He found high correlations be-

tween lay and professional judges using a direct magnitude estimation procedure for scaling severity and concluded that the procedure was satisfactory for scaling stuttering.

### **Adaptation and Stuttering**

Neelley and Timmons (1967) studied fluency adaptation in stutterers and nonstutterers and found that adaptation occurred for both groups but was more pronounced in the former group. Wingate (1966b) found that altering prosody affected adaptation during sequential readings. He interpreted his findings as indicating that observed adaptation in stuttering is due to organismic adaptation to situations and increasing familiarity with reading material. Siegel and Haugen (1964) found that varying audience size affected adaptation and frequency of stuttering. As audience size increased, adaptation decreased. Initial stuttering frequency was inversely related to adaptation rate. Sheehan, Hadley, and Gould (1967) reported that adaptation was less when Ss read to a peer group than when they read to an authority group. Severity was greater in the presence of the latter group.

### **Linguistic Variables and Stuttering**

Bloodstein and Gantwerk (1967) investigated relationships between grammatical function of words and stuttering in young children. They observed that a true grammatical factor does not exist during the initial phase of stuttering. Knabe, Nelson, and Williams (1966) did a comparative analysis of linguistic output of adult stutterers and nonstutterers. They concluded that groups were similar on all variables studied.

Soderberg (1966) looked at relationships among stuttering, word length, and word frequency. Results indicated that stuttering was directly related to word length and inversely related to word frequency, with the former being the more powerful variable. Wingate (1967) found more stuttering on two-syllable words than on one-syllable words.

Quarrington (1965) investigated stuttering as a function of the information value and sentence position of words. Stuttering was found to increase on high information words and on words occurring early in sentences. The latter was considered to be the more powerful variable. Schlesinger and others (1965) analyzed recordings of stutterers and found that words with high transition probability were stuttered on half as often as words with low transition probability.

### **Characteristics of Stutterers**

Gildston (1967) used the Hilden Q Sort with stutterers and a control group. Findings suggested that stutterers were less self-accepting and per-



ceived their parents as accepting them less. Wingate (1966a) studied behavioral rigidity and found the stutterers to be less flexible than the control group in mental tests requiring rapid and contiguous changes in set. Moral rigidity did not differ between groups.

Adams and Dietze (1965) showed that the latency period for stutterers reacting to a word association test was longer than that for nonstutterers and that the period was longest for words pertaining to guilt. Shearer and Simmons (1965) used an acoustic impedance bridge to study middle ear muscle activity during speech. No differences were found between stutterers and nonstutterers.

### Modification of Stuttering

Martin and Siegel (1966a) found that response-contingent electric shock reduced stuttering frequency to zero while removal of shock resulted in a return of stuttering to the base rate. Using shock, they independently manipulated specific stuttering behaviors. Stuttering was brought under the control of a discriminative stimulus. In a subsequent study, these authors (1966b) punished stuttering and rewarded fluency. Again stuttering was reduced and brought under the control of a discriminative stimulus. Quist and Martin (1967) found that presenting the word *wrong* contingent upon stuttering reduced its frequency. An extensive discussion of research related to operant procedures and stuttering was presented in a book edited by Sloane and MacAulay (1968).

Sheehan and Martyn (1966) reviewed data obtained from Ss who indicated that they had been stutterers but had spontaneously recovered. The authors concluded that these Ss were never severe stutterers, that in stuttering they repeated syllables rather than sounds or words, and that their familial patterns were more similar to nonstutterers than to stutterers. Shearer and Williams (1965) studied a group similar to Sheehan's. Their Ss reported that speaking more slowly was one of the things they found most beneficial in handling their problem.

Lanyon (1965) investigated the hypothesis that high adaptation and low consistency scores indicate a good prognosis for therapy. Findings were somewhat positive, but the author indicated a need for measures with less variability. Lanyon (1966) used the *Minnesota Multiphasic Personality Inventory (MMPI)* to predict therapy results and found ego strength scale scores to be positively correlated with therapy success, and deviant response scale scores to be negatively correlated.

### Voice Disorders

During the past few years, voice scientists have made significant progress in developing an understanding of vocal physiology and acoustics. However, a minimal amount of work has been done on clinical problems.

## Analysis of Voice Disorders

Shipp and Huntington (1965) studied laryngitic and postlaryngitic perceptual factors in acute laryngitic hoarseness. They found hoarseness related to the perception of breathiness, voice aberrations, and restricted frequency range. Lower pitch and harshness did not characterize acute laryngitic hoarseness. Michel and Hollien (1968) used trained and untrained judges to obtain psychophysical measures on Ss with harsh voices and on Ss producing vocal fry. They found that clinical harshness and vocal fry are distinct perceptual entities.

Yanagihara (1967) used acoustic analysis techniques and synthesis by means of spectrography to study hoarseness. He found the main determiners of hoarseness to be interactions of noise components in the main format of high frequency noise components and loss of high frequency harmonic components. As hoarseness increased these factors became more prominent. Using these factors, Yanagihara identified four types of hoarseness. Moore and Thompson (1965) used high-speed laryngeal photography and phonelagrams to study hoarse voice. They found that random fundamental frequency variations exist in hoarse voices. Both techniques used in the study could independently detect these variations.

## Identification and Evaluation of Voice Disorders

Davis and Boone (1967) found no differences between normal Ss and Ss with hyperfunctional voice disorders on the pitch discrimination and tonal memory subtests of the *Seashore Measures of Musical Talents*. Aronson, Peterson, and Litin (1966) used clinical psychiatric procedures and the *MMPI* to study Ss with functional voice disorders. They found no serious psychopathology, but reported that, when subjects were evaluated in terms of current definitions of conversion and hysteria, almost all met criteria for the former and less than half for the latter. *MMPI* data tended to support psychiatric data.

Baynes (1966) surveyed first-, third-, and sixth-grade students in a public school and found chronic hoarseness in 7.1 percent, with the highest prevalence in first-grade students. Schwartz, Noyek, and Nolberg (1966) studied 1,000 consecutive patients with persistent hoarseness. They found carcinoma in 19.6 percent and premalignant lesions in another 20 percent. Polyps or vocal nodules were found in 39 percent and laryngitis in 16 percent. Other factors found were chord paralysis, laryngitis sicca, tubercular laryngitis, adductor insufficiency, and psychogenic hoarseness.

## Treatment of Voice Disorders

Von Leden, Yanagihara, and Werner-Kukuk (1967) reported on a series of Ss with unilateral vocal chord paralysis who were treated with Teflon injections for their paralyzed chord. According to the authors, aerodynamic

and acoustic measures and analysis of high-speed laryngeal photography indicated marked improvement in physiologic and acoustic parameters.

## Delayed Speech and Language

Linguists, psychologists, and others have devoted extensive time and effort to develop an understanding of the processes by which normal language is acquired. Their efforts may eventually aid in the management of delayed speech and language. Reports of research directly related to clinical management problems have been sparse. Major research concerns have been the identification of perceptual factors which might account for delay and the application of operant conditioning procedures to develop speech in mute children.

### Perceptual Factors Associated with Delayed Speech and Language

Stark (1967) studied sequencing abilities of aphasic children using the *Knox Cube Test*, a tapping test, and auditory, vocal, visual, and motor sequencing subtests from the *ITPA*. Ss were below norms on all tests, with their lowest scores on auditory vocal tests. The author interpreted his findings as supporting the concept that temporal sequencing is the core of the aphasic child's problems. In another study, Stark, Poppen, and May (1967) looked at the effects of alterations of prosodic features on sequencing performance of aphasic children. Results indicated that sequencing difficulties were due to forgetting the first items in a series. When first items were stressed, recall was enhanced. Similar effects were not found in normals. Aten and Davis (1968) investigated disturbances in perceptions of auditory sequences in children with minimal cerebral dysfunction. Their experimental group scored lower than a control group of normals on non-verbal tests, backward digit span, serial noun span, multisyllabic word repetition, scrambled sentence arrangements, and oral sequential accuracy. The authors concluded that temporal sequencing and reduced attention are basic to children's language problems.

### Diagnosis and Therapy

McReynolds (1966) utilized operant conditioning procedures to test speech sound discrimination in aphasic and normal children. Findings indicated that aphasics discriminated as well as normals between sounds in isolation but were inferior to normals when sounds were placed in words.

Hingtgen and Coulter (1967) used operant conditioning techniques with mute autistic children to establish associations between several pairs of auditory stimuli, lever pressing, and receipt of food. They interpreted their results as indicating that some mute autistics are capable of learning simple auditory-motor associations.



Cowan, Hoddinot, and Wright (1965) studied autistic children who responded inconsistently to verbal requests to perform tasks. They subjected these children to conditioning trials in which correct responses were rewarded. Some Ss attained consistent responses for one task and were then able to perform the other. Other Ss remained inconsistent. The authors interpreted their data as indicating a need to distinguish between competency and willingness to perform tasks. Hewett (1965); Wolf, Risley, and Mees (1964); and Lovaas and others (1966) reported successful use of operant conditioning procedures in establishing verbal skills in mute children.

## Aphasia

Research on aphasia and related language disturbances has reflected a strong interest in linguistic analysis and behavioral models.

### Language, Cognition, and Perception Studies

Pizzamiglio and Black (1968) analyzed errors on predictive responses of aphasic adults performing a sentence-completion and picture task in which Ss typed one-word responses. Results indicated that incorrect predictions related to frequency of occurrence of words in English, and the vowels could be predicted to occur in appropriate places in words. The authors suggested that errors were due to perseveration and anticipation of letters in the final portions of words. Noll and Hoops (1967) looked at spelling errors of mild aphasics and found that nonpropositional morphemes and words presented difficulties and that errors were related to length of word. No relationship was found between spelling and educational level. Jenkins and Schuell (1965) used tests of sentence comprehension and detection of errors with normal and aphasic Ss. Aphasics showed significantly more acquiescence response bias on all tests.

Halpern (1965) studied the effects of stimulus variables on dysphasics' verbal errors. Analysis of data indicated that errors varied in relation to abstraction level when stimuli were presented in the visual and the visual-auditory modalities. Regardless of modality, verbs and adjectives produced more errors than nouns, long words produced more errors than short words, and frequently occurring words produced more errors than infrequently occurring words. In general, visual stimuli presentations produced the most errors. Goodglass, Fodor, and Schulhoff (1967) studied conditions under which function words of grammar are confused or omitted in the speech of aphasics. They found that prosodic characteristics of function words determined whether they are lost or retained in agrammatic speech.

Goodglass and others (1966) looked at specific semantic work categories in the speech of aphasics. Tests used included naming of objects, colors, numbers, letters, and actions, and auditory comprehension of words in naming and other categories. They concluded that patterns of differences

in naming and auditory discrimination among words of various semantic categories were predictive in relation to the major clinical aphasia types. Sefer and Henrikson (1966) used a word association test with aphasic and nonaphasic Ss. They found that aphasics made fewer homogeneous responses than nonaphasics, that aphasics' word associations followed the same pattern of variance by parts of speech as nonaphasics, and that syntagmatic and paradigmatic responses were negatively correlated for both groups.

Archibald, Wepman, and Jones (1967) investigated nonverbal cognitive performance of aphasic and nonaphasic brain-damaged adults. They found that more cognitive impairments occurred among severe aphasics and that the Elithorn Mazes differentiated right from left brain-damaged Ss. They concluded that aphasia is specifically a defect of language and memory for language.

Sigenthaler and Goldstein (1967) studied auditory and visual figure-background problems in adult aphasics. Both problems were found to be present but appeared to vary in severity independently. Ebbin and Edwards (1967) matched aphasic and nonaphasic brain-damaged adults and tested their speech sound discrimination at different intersound intervals. Aphasics made more errors with short intervals, and their overall performance was found to be related to comprehension ability but not to auditory recognition or retention skills.

### **Lesion Site and Speech and Language Behavior**

Benson (1967) found (a) low verbal output, dysprosody, dysarthria, and effort in speech production to be associated with brain lesions anterior to the Fissure of Rolando, and (b) paraphasia, press of speech, and lack of substantive words to be associated with the more posterior lesions. De Renzi and Spinnler (1967) looked at performance of left brain-damaged Ss (aphasics and nonaphasics), right brain-damaged Ss, and normal subjects. They found that right brain-damaged Ss scored lower than left brain-damaged Ss on purely perceptual tasks and that aphasics were more impaired on tests requiring language.

Bonkowski (1967) presented a sample word and a set of words in a unimodal (visual or auditory) and in a cross-modal (some visual and some auditory) manner to Ss with left or right brain damage. Ss responded by matching, spelling, or choosing a synonym from the set of words. Results indicated that Ss with right brain damage responded slowly and made more errors on items with extraverbal components. The converse occurred on items requiring facility with verbal language components. Modality differences were not significant. De Renzi, Pieczuro, and Vignolo (1966) looked for oral apraxia in left and right brain-damaged Ss and in normal Ss. Right brain-damaged Ss and normal Ss scored similarly on apraxia tests. Apraxia was found to be common in left brain-damaged Ss with Broca's aphasia, phonemic jargon aphasia, and conduction aphasia.

## Diagnosis and Therapy for Aphasia

Spiegel, Jones, and Wepman (1965) used the *Language Modalities Test for Aphasia* to predict free speech characteristics of aphasics. They concluded that they could predict some general language features. They interpreted their findings as being in conflict with a unidimensional concept of aphasics' language disorder.

Edwards (1965) used an automated procedure to train aphasics in a "matching-to-sample" task. He concluded that it seems possible to teach aphasics, previously thought to be untrainable, any program which can utilize visual discrimination. Rosenberg (1965) studied aphasics' performance on automated visuo-perceptual discrimination training and transfer tasks. His data were interpreted as supporting the view that form discrimination is a basic factor in visual-verbal aspects of language behavior and that it is possible to design automated procedures for training aphasics. An extensive discussion of research relating to diagnosis and therapy was presented in a book by Schuell, Jenkins, and Jiménez-Pabón (1964).

## Bibliography

- ADAMS, MARTIN R., and DIETZE, DORIS A. "A Comparison of Reaction Times of Stutterers and Nonstutterers to Items on a Word Association Test." *Journal of Speech and Hearing Research* 8: 195-202; June 1965.
- ARCHIBALD, Y. M.; WEPMAN, J. M.; and JONES, L. V. "Nonverbal Cognitive Performance in Aphasic and Nonaphasic Brain-Damaged Patients." *Cortex* 3: 276-94; September 1967.
- ARONSON, ARNOLD; PETERSON, HERBERT W., JR.; and LITIN, EDWARD M. "Psychiatric Symptomatology in Functional Dysphonia and Aphonia." *Journal of Speech and Hearing Disorders* 31: 115-27; May 1966.
- ATEN, JAMES, and DAVIS, JAMES. "Disturbances in the Perception of Auditory Sequence in Children with Minimal Cerebral Dysfunction." *Journal of Speech and Hearing Research* 11: 236-45; June 1968.
- BAER, WILLIAM P., and WINITZ, HARRIS. "Acquisition of /v/ in 'Words' as a Function of the Consistency of /v/ Errors." *Journal of Speech and Hearing Research* 11: 316-33; June 1968.
- BAYNES, ROBERT A. "An Incidence Study of Chronic Hoarseness Among Children." *Journal of Speech and Hearing Disorders* 31: 172-75; May 1966.
- BENSON, D. FRANK. "Fluency in Aphasia: Correlation with Radioactive Scan Localization." *Cortex* 3: 373-94; December 1967.
- BLOODSTEIN, OLIVER, and GANTWERK, BARBARA F. "Grammatical Function in Relation to Stuttering in Young Children." *Journal of Speech and Hearing Research* 10: 786-89; December 1967.
- BONKOWSKI, ROBERT J. "Verbal and Extraverbal Components of Language as Related to Lateralized Brain Damage." *Journal of Speech and Hearing Research* 10: 558-64; September 1967.
- BOSMA, JAMES F., editor. *Symposium on Oral Sensation and Perception*. Springfield, Ill.: Charles C. Thomas, 1967. 300 pp.
- CHRISTINE, DOROTHY, and CHRISTINE, CHARLES. "The Relationship of Auditory Discrimination to Articulatory Defects and Reading Retardation." *Elementary School Journal* 65: 97-100; November 1964.
- COWAN, PHILLIP A.; HODDINOT, B. A.; and WRIGHT, BARBARA A. "Compliance and Resistance in the Conditioning of Autistic Children: An Exploratory Study." *Child Development* 36: 913-23; December 1965.



- DAVIS, DEMARIS S., and BOONE, DANIEL R. "Pitch Discrimination and Tonal Memory Abilities in Adult Voice Patients." *Journal of Speech and Hearing Research* 10: 811-15; December 1967.
- DE RENZI, E., and SPINNIER, H. "Impaired Performance on Color Tasks in Patients with Hemisphere Damage." *Cortex* 3: 194-217; April 1967.
- DE RENZI, E.; PIECZURO, A.; and VIGNOLO, L. A. "Oral Apraxia and Aphasia." *Cortex* 2: 50-73; January 1966.
- EBBIN, JUDITH B., and EDWARDS, ALLAN E. "Speech Sound Discrimination of Aphasics when Intersound Interval is Varied." *Journal of Speech and Hearing Research* 10: 120-25; March 1967.
- EDWARDS, ALLAN E., "Automated Training for a 'Matching-to-Sample' Task in Aphasia." *Journal of Speech and Hearing Research* 8: 39-42; March 1965.
- ELBERT, MARY; SHELTON, RALPH L.; and ARNDT, WILLIAM B., JR. "A Task for Evaluation of Articulation Change: I. Development of Methodology." *Journal of Speech and Hearing Research* 10: 281-88; June 1967.
- FLETCHER, SAMUEL G., and MELDRUM, JAMES R. "Lingual Function and Relative Length of Lingual Frenulum." *Journal of Speech and Hearing Research* 11: 382-90; June 1968.
- GILDSTON, PHYLLIS. "Stutterers' Self-Acceptance and Perceived Parental Acceptance." *Journal of Abnormal Psychology* 72: 59-64; February 1967.
- GILMORE, STUART I., and FAMILANT, ROSALEE P. "Race, Residence, Socioeconomic Status, and Responses to Articulation Test Stimuli." *Southern Speech Journal* 33: 44-49; Fall 1967.
- GOODCLASS, HAROLD; FODOR, IRIS G.; and SCHULHOFF, CELIA. "Prosodic Factors in Grammar—Evidence from Aphasia." *Journal of Speech and Hearing Research* 10: 5-20; March 1967.
- GOODCLASS, HAROLD, and OTHERS. "Specific Semantic Work Categories in Aphasia." *Cortex* 2: 74-89; January 1966.
- GRIFFITH, JERRY. "The Modification of Functional Articulation Errors Under Principles of Instrumental Conditioning." *Speech Monographs* 32: 178-84; June 1965.
- HALLER, RAPHAEL M. "Effect of Background Speech on Phoneme Discrimination Abilities of Children with Normal and Defective Articulation." *Journal of Auditory Research* 7: 77-84; January 1967.
- HALPERN, HARVEY. "Effect of Stimulus Variables on Dysphasic Verbal Errors." *Perceptual and Motor Skills* 21: 291-98; August 1965.
- HEIGHTON, ROBERT S., and NICHOLS, ALAN C. "Effects of Untreated 'Open Bite' Malocclusion on the Articulation of Children." *Journal of Communication Disorders* 1: 166-69; December 1967.
- HEWETT, FRANK M. "Teaching Speech to an Autistic Child Through Operant Conditioning." *American Journal of Orthopsychiatry* 35: 926-36; October 1965.
- HINGTGEN, JOSEPH, and COULTER, SUSAN K. "Auditory Control of Operant Behavior in Mute Autistic Children." *Perceptual and Motor Skills* 25: 561-65; August 1967.
- HUTCHINSON, BARBARA B., and CLARK, DOROTHY W. "Auditory and Visual Discrimination Skills of Normal and Articulation-Defective Children." *Perceptual and Motor Skills* 26: 259-65; February 1968.
- JENKINS, EDNA, and LOHR, FRANCES E. "Severe Articulation Disorders and Motor Ability." *Journal of Speech and Hearing Disorders* 29: 286-92; August 1964.
- JENKINS, JAMES, and SCHUELL, HILDRED. "Acquiescence Response Set in Aphasics." *Journal of Abnormal Psychology* 70: 111-13; April 1965.
- KNABE, JUDITH M.; NELSON, LOIS A.; and WILLIAMS, FREDERICK. "Some General Characteristics of Linguistic Output: Stutterers Versus Nonstutterers." *Journal of Speech and Hearing Disorders* 31: 178-82; May 1966.
- KORST, JOSEPH W. "A Comparison of Results from the Peabody Picture Vocabulary Test and Leiter International Performance Scale with Children Having Functional Articulatory Disorders." *Cerebral Palsy Journal* 27: 3-5; September 1966.
- LANYON, RICHARD I. "The Relationship of Adaptation and Consistency to Improvement in Stuttering Therapy." *Journal of Speech and Hearing Research* 8: 263-69; September 1965.
- LANYON, RICHARD I. "The MMPI and Prognosis in Stuttering Therapy." *Journal of Speech and Hearing Disorders* 31: 186-91; May 1966.

- LICHTENBERG, FRANCIS S. "A Comparison of Children's Ability To Make Speech Sound Discriminations." *Volta Review* 68: 426-34; June 1966.
- LOVAAS, O. IVAR, and OTHERS. "Acquisition of Imitative Speech by Schizophrenic Children." *Science* 151: 705-707; February 1966.
- MARTIN, RICHARD. "Direct Magnitude-Estimation Judgments of Stuttering Severity Using Audible and Audible-Visible Speech Samples." *Speech Monographs* 32: 169-77; June 1965.
- MARTIN, RICHARD R., and SIEGEL, GERALD M. "The Effects of Response Contingent Shock on Stuttering." *Journal of Speech and Hearing Research* 9: 340-52; September 1966. (a)
- MARTIN, RICHARD R., and SIEGEL, GERALD M. "The Effects of Simultaneously Punishing Stuttering and Rewarding Fluency." *Journal of Speech and Hearing Research* 9: 466-75; September 1966 (b)
- McREYNOLDS, LELIA V. "Operant Conditioning for Investigating Speech Sound Discrimination in Aphasic Children." *Journal of Speech and Hearing Research* 9: 519-28; December 1966.
- MICHEL, JOHN F., and HOLLIEN, HARRY. "Perceptual Differentiation of Vocal Fry and Harshness." *Journal of Speech and Hearing Research* 11: 439-43; June 1968.
- MOORE, PAUL, and THOMPSON, CARL L. "Comments on Physiology of Hoarseness." *Archives of Otolaryngology* 81: 97-102; January 1965.
- NEELLEY, JAMES N., and TIMMONS, ROY J. "Adaptation and Consistency in the Disfluent Speech Behavior of Young Stutterers and Nonstutterers." *Journal of Speech and Hearing Research* 10: 250-56; June 1967.
- NICHOLS, ALAN C. "Utility Factors in Articulation Screening: Efficiency of Selection and Cost." *Journal of Communication Disorders* 1: 16-25; May 1967.
- NOLL, J. DOUGLAS, and HOOPS, H. RAY. "Aphasic Grammatical Involvement as Indicated by Spelling Ability." *Cortex* 3: 419-32; December 1967.
- PALMER, MARTIN F.; WOURTH, CHARLES W.; and KINCHELOE, JOHN W. "The Incidence of Lingual Apraxia and Agnosia in 'Functional' Disorders of Articulation." *Cerebral Palsy Review* 25: 7-9; September 1964.
- PIZZAMIGLIO, LUIGI, and BLACK, JOHN W. "Phonic Trends in the Writing of Aphasic Patients." *Journal of Speech and Hearing Research* 11: 77-84; March 1968.
- PRINS, DAVID, and BLOOMER, H. HARLAN. "Consonant Intelligibility: A Procedure for Evaluating Speech in Oral Cleft Subjects." *Journal of Speech and Hearing Research* 11: 128-37; March 1968.
- QUARRINGTON, BRUCE. "Stuttering as a Function of the Information Value and Sentence Position of Words." *Journal of Abnormal Psychology* 70: 221-24; June 1965.
- QUIST, RAYMOND W., and MARTIN, RICHARD R. "The Effect of Response Contingent Verbal Punishment on Stuttering." *Journal of Speech and Hearing Research* 10: 795-800; December 1967.
- RECHNER, JOAN, and WILSON, BETTY A. "Relation of Speech Sound Discrimination and Selected Language Skills." *Journal of Communication Disorders* 1: 26-30; May 1967.
- RINGEL, ROBERT L., and EWANOWSKI, STANLEY. "Oral Perception: I. Two-Point Discrimination." *Journal of Speech and Hearing Research* 8: 389-98; December 1965.
- RINGEL, ROBERT L., and FLETCHER, HILARY M. "Oral Perception: III. Texture Discrimination." *Journal of Speech and Hearing Research* 10: 642-49; September 1967.
- RINGEL, ROBERT L.; SAXMAN, JOHN H.; and BROOKS, ALTA R. "Oral Perception: II. Mandibular Kinesthesia." *Journal of Speech and Hearing Research* 10: 637-41; September 1967.
- RONSON, IRWIN. "Incidence of Visceral Swallow Among Lispers." *Journal of Speech and Hearing Disorders* 30: 318-24; November 1965.
- ROSENBERG, BETH. "The Performance of Aphasics on Automated Visuo-Perceptual Discrimination Training, and Transfer Tasks." *Journal of Speech and Hearing Research* 8: 165-81; June 1965.
- SCHLESINGER, I. M., and OTHERS. "Stuttering Information Load, and Response Strength." *Journal of Speech and Hearing Disorders* 30: 32-36; February 1965.
- SCHUELL, HILDRED; JENKINS, JAMES; and JIMÉNEZ-PABÓN, EDWARD. *Aphasia in Adults*. New York: Harper and Row, 1964. 428 pp.

- SCHWARTZ, LEO; NOYK, ARNOLD M.; and NOLBERG, DAVID. "Persistent Hoarseness: Etiology in 1000 Consecutive Patients." *New York State Journal of Medicine* 66: 2658-62; October 1966.
- SEFER, JOYCE W., and HENRIKSON, ERNEST H. "The Relationship Between Word Association and Grammatical Classes in Aphasia." *Journal of Speech and Hearing Research* 9: 529-41; December 1966.
- SHEARER, WILLIAM M., and SIMMONS, F. BLAIR. "Middle Ear Activity During Speech in Normal Speakers and Stutterers." *Journal of Speech and Hearing Research* 8: 203-207; June 1965.
- SHEARER, WILLIAM M., and WILLIAMS, J. DAVID. "Self-Recovery from Stuttering." *Journal of Speech and Hearing Disorders* 30: 288-90; August 1965.
- SHEEHAN, JOSEPH G., and MARTYN, MARGARET M. "Spontaneous Recovery from Stuttering." *Journal of Speech and Hearing Research* 9: 121-35; March 1966.
- SHEEHAN, JOSEPH G.; HADLEY, ROBERT; and GOULD, EDWARD. "Impact of Authority on Stuttering." *Journal of Abnormal Psychology* 72: 290-93; June 1967.
- SHELTON, RALPH L., and OTHERS. "Identification of Persons with Articulation Errors from Observation of Non-Speech Movements." *American Journal of Physical Medicine* 45: 143-50; June 1966.
- SHELTON, RALPH L.; ELBERT, MARY; and ARNDT, WILLIAM B., JR. "A Task for Evaluation of Articulation Change: II. Comparison of Task Scores During Baseline and Lesson Series Testing." *Journal of Speech and Hearing Research* 10: 578-85; September 1967.
- SHERMAN, DOROTHY, and GEITH, ANNETTE. "Speech Sound Discrimination and Articulation Skill." *Journal of Speech and Hearing Research* 10: 277-80; June 1967.
- SHIPP, THOMAS, and HUNTINGTON, DOROTHY A. "Some Acoustic and Perceptual Factors in Acute-Laryngitic Hoarseness." *Journal of Speech and Hearing Disorders* 30: 350-59; November 1965.
- SIEGEL, GERALD M., and HAUGEN, DAVID. "Audience Size and Variations in Stuttering Behavior." *Journal of Speech and Hearing Research* 7: 381-88; December 1964.
- SIGENTHALER, BRUCE M., and GOLDSTEIN, JEROD. "Auditory and Visual Figure-Background Perception by Adult Aphasics." *Journal of Communication Disorders* 1: 152-58; December 1967.
- SLOANE, HOWARD N., JR., and MACAULAY, BARBARA D. *Operant Procedures in Remedial Speech and Language Training*. Boston: Houghton Mifflin Co., 1968. 444 pp.
- SMITH, MYRNA W., and AINSWORTH, STANLEY S. "The Effects of Three Types of Stimulation on Articulatory Responses of Speech Defective Children." *Journal of Speech and Hearing Research* 10: 333-38; June 1967.
- SNIDECOR, JOHN C., and KARRIS, ANTHONY K. "A Speech Corrective Prosthesis for Anterior 'Open-Bite' Malocclusion and Its Effects on Two Postdental Fricative Sounds." *Journal of Prosthetic Dentistry* 15: 779-84; July 1965.
- SODERBERG, GEORGE A. "The Relations of Stuttering to Word Length and Word Frequency." *Journal of Speech and Hearing Research* 9: 584-89; December 1966.
- SOMMERS, RONALD K., and OTHERS. "The Effectiveness of Group and Individual Therapy." *Journal of Speech and Hearing Research* 9: 219-25; June 1966.
- SOMMERS, RONALD K., and OTHERS. "Factors Related to the Effectiveness of Articulation Therapy for Kindergarten, First, and Second Grade Children." *Journal of Speech and Hearing Research* 10: 428-37; September 1967.
- SPIEGEL, DOUGLAS K.; JONES, LYLE V.; and WEPMAN, JOSEPH M. "Test Responses as Predictors of Free-Speech Characteristics in Aphasia Patients." *Journal of Speech and Hearing Research* 8: 349-62; December 1965.
- STARK, JOEL. "A Comparison of the Performance of Aphasic Children on Three Sequencing Tests." *Journal of Communication Disorders* 1: 31-34; May 1967.
- STARK, JOEL; POPPEN, ROGER; and MAY, MICHAEL Z. "Effects of Alterations of Prosodic Features on the Sequencing Performance of Aphasic Children." *Journal of Speech and Hearing Research* 10: 849-55; December 1967.
- STOIA, LOUIS, and REELING, GLENN E. "Better Speech for Head Start Children." *Elementary School Journal* 67: 213-17; January 1967.
- SUTELNY, JOANNE D.; MESTRE, JORGE C.; and SUTELNY, J. DANIEL. "Comparative Study of Normal and Defective Articulation of /s/ as Related to Malocclusion and Deglutition." *Journal of Speech and Hearing Disorders* 29: 269-85; August 1964.



- VON LEDEN, HANS; YANAGIHARA, NAOAKI; and WERNER-KUKUK, ELVIRA. "Teflon in Unilateral Vocal Chord Paralysis." *Archives of Otolaryngology* 85: 666-74; June 1967.
- WEINBERG, BERND. "A Cephalometric Study of Normal and Defective /s/ Articulation and Variations in Incisor Dentition." *Journal of Speech and Hearing Research* 11: 288-300; June 1968.
- WEINER, PAUL S. "Auditory Discrimination and Articulation." *Journal of Speech and Hearing Disorders* 32: 19-28; February 1967.
- WILSON, FRANK B. "Efficacy of Speech Therapy with Educable Mentally Retarded Children." *Journal of Speech and Hearing Research* 9: 423-33; September 1966.
- WINGATE, M. E. "Behavioral Rigidity in Stutterers." *Journal of Speech and Hearing Research* 9: 626-29; December 1966. (a)
- WINGATE, M. E. "Prosody in Stuttering Adaptation." *Journal of Speech and Hearing Research* 9: 550-56; December 1966. (b)
- WINGATE, M. E. "Stuttering and Word Length." *Journal of Speech and Hearing Research* 10: 146-52; March 1967.
- WINITZ, HARRIS, and PREISLER, LINDA. "Effect of Distinctive Feature Pretraining in Phoneme Discrimination Learning." *Journal of Speech and Hearing Research* 10: 515-30; September 1967.
- WOLF, MONTROSE; RISLEY, TODD; and MEES, HAYDEN. "Application of Operant Conditioning Procedures to the Behavior Problems of an Autistic Child." *Behavior Research and Therapy* 1: 305-12; March 1964.
- YANAGIHARA, NAOAKI. "Significance of Harmonic Changes and Noise Components in Hoarseness." *Journal of Speech and Hearing Research* 10: 531-41; September 1967.

## CHAPTER IV

### The Visually Handicapped

CARSON Y. NOLAN and SAMUEL C. ASHCROFT\*

Current research has followed the trends implied in the last REVIEW (Ashcroft and Harley, 1966) devoted to this topic. While technological research continued to receive emphasis, interest in research on tests and appraisal, abstract functioning, and concept development appeared to be increasing. At long last, the area of mobility was subject to some basic research, and studies of the multiple-handicapped blind proliferated.

The first annual report of the National Advisory Committee on Handicapped Children (1968) recommended the promotion of research and development centers that will focus a sustained effort on major research problems in the education of handicapped children. While there has been an encouraging increase in research projects and reports on them, the period under review reflects too little sustained attention to problems of continuing concern. Research involving intervention in education continues to be scarce. Research in the area of the visually handicapped tends to follow rather than to initiate practice. Implementation of research findings is slow to be accomplished. There is some evidence that findings of research on the use of residual vision and learning by listening were being applied during the review period.

#### General Aspects

The major dependable sources of research and research information continued to be the American Printing House for the Blind (APH), American Foundation for the Blind (AFB), and Massachusetts Institute of Technology (MIT). With the possible exception of MIT, where there has been concentration on technological research, no university has yet emerged providing the "critical mass" of needed research interests and personnel. The Center for Rate Controlled Recordings at the University of Louisville is a promising development, but it too has a very limited focus (Foulke, 1967). The U.S. Office of Education has been the source of useful statistics, and its new intramural research authority may enable the expansion of such efforts and the invitation of project research. USOE has yet to support a research program—a needed development in this area of emphasis.

\* The authors would like to express appreciation to those graduate students at Peabody College, especially Roy Brothers, and to June Morris and Rita Wright at the American Printing House for the Blind, who helped in the preparation of this manuscript.

In the USOE, Jones and Collins (1966) analyzed reports from 353 special local public school programs employing one or more full-time teachers of visually handicapped children from 54 residential schools. Their brief report provided useful data on (a) the number and percent of special residential and local school programs by type of visual handicap served, (b) organizational patterns for educational services (itinerant most common), (c) procedures utilized for determining mode of reading, and (d) prevalence of a visual handicap in school-aged children (1:1,000). The Jones and Collins study indicated that by 1964 more than 50 percent of all pupils in the nation's public schools had access to full-time special teachers of the visually handicapped in their home or neighboring communities. The Southeast continued to lag in the development of such local programs.

The widespread change from special education approaches emphasizing traditional categories to learning characteristics was not reflected in research or in general works produced during the past three years. Texts in special education, while generally eschewing traditional treatment of the education of visually handicapped children, usually failed to consider more than inferentially learning disability applications in the emphasis on the visually handicapped. Bateman (1967) provided the outstanding chapter, an excellent work integrating, interpreting, and applying research to educational programs and problems. Although it too fell short in the learning disabilities aspect, it was in marked contrast to other chapters in general special education texts. The chapter, more adequately than any before, reviewed research on partially seeing children and pointedly raised the question, "Is special education (beyond the provision of large-type books, etc.) necessary or beneficial for the partially seeing child?"

Johnson and Myklebust (1967) indicated that learning disabilities are of many types, but "homogeneity is derived from the fact that it is the neurology of learning that has been disrupted." They felt that the neurology of learning is not disturbed in the deaf or the blind child. Lawson (1968) provided an excellent review of ophthalmological factors in reading and other aspects of growth, development, and education. In his summary, he stated that the role of ocular factors in the cause of reading retardation has been refuted, but eye defects "can aggravate a learning disorder or impede response to educational remediation." Included in his chapter was a report on a study of 82 children with learning disabilities, 46 percent of whom had positive eye findings including hyperopia, astigmatism, strabismus, myopia, crossed dominance, and anisometropia. The positive and negative groups differed on only 2 out of 19 behavioral variables. When grouped by use of a multiple discriminant analysis technique, eight variables (primarily motor abilities) distinguished between the groups. The investigator cautioned that the findings were preliminary in nature.

Birch and others (1966) reported one of the few studies on partially seeing children. Their study appeared to raise serious questions regarding



the efficacy of special programs. While the average sixth-grade partially seeing pupil was of at least average intelligence and had had some type of special program for  $5\frac{1}{2}$  of his 8 years of school, he was overage for grade and  $2\frac{1}{2}$  years retarded in academic achievement. The investigators called attention to sampling and other problems in the study and noted that comparison with the most recent information on this point (Jones and Collins, 1966) could not be made because the Jones and Collins data were in terms of numbers of programs and teachers and their data were in terms of numbers of pupils. Nevertheless, as Bateman has stated, the effectiveness of educational programs for the partially seeing is questionable. Birch and his co-investigators claimed that the most serious educational problem demonstrated by partially seeing children is severe educational retardation with respect to capacity for achievement, a condition which "appears not to be related in any substantial way to types of visual disabilities categorized by a structural classification or to degree of visual acuity."

Nolan (1967) noted that between 1963 and 1966 there was a 50 percent increase in the proportion of children classified as ungraded and that local school enrollments grew 15 percent while residential schools remained static. The number of legally blind children registered with the American Printing House continued to increase. In 1968, this registration totalled 20,266, with 59 percent of these children being enrolled in public school programs.

### Technological Research

Continuing interest in technological research was focused on displays for tactile reading, systems for braille production, and devices to enhance mobility.

Linville and Bliss (1966) reported on the development of a direct-translation reading aid for the blind. Their device produces a tactual facsimile of ordinary print which is displayed by a dense array of pins which can be made to vibrate individually through perforations in a plate. Using this device, blind subjects have attained reading rates of about 30 words per minute.

Progress on development of a braille reader which provides a continuous display of braille in standard form from a punched paper-tape input is reported in the *Proceedings—Conference on New Processes for Braille Manufacture* sponsored by the Sensory Aids Evaluation and Development Center (1968). Also reported in these *Proceedings* are developments in computer translation of braille, the adaptation of compositors tapes for braille translation, and the development of an electric braille typewriter.

The state of the development of mobility aids was well documented in the *Proceedings of the Rotterdam Mobility Research Conference* published by the American Foundation for the Blind (1965). The status of a variety of environmental probes was reported, together with descriptions of trial uses of many of these devices.

A frequently ignored yet persistent question is the economic feasibility of the application of the outcomes of much technological research. Researchers in technological fields continue to ignore this facet of the problem with the result that practical applications of their research have been few.

## Psychological Aspects

### Social Acceptance

The social acceptance of exceptional children at the high school level was the topic of a report by Jones, Gottfried, and Owens (1966). The authors required 186 high school students to respond to a questionnaire on which certain groups of exceptional and average persons were paired in all possible combinations. The subjects were required to choose the one group of the pair which they preferred with regard to a particular statement. There were marked differences among responses to different exceptionalities in a variety of interpersonal situations. In most situations, the average and gifted were most acceptable while the severely mentally retarded were least acceptable. Among the 12 groups from which subjects were to choose their preference, blind individuals ranked from sixth to eleventh depending on the interpersonal situation. Therefore they were among the least well accepted. These results provide some evidence of the continued existence of negative stereotypes for the blind.

### Tests and Appraisal

The usual heavy emphasis on psychological and educational tests is reflected in the research reviewed for this reporting period. The performance of the blind and sighted children on the *Wechsler Intelligence Scale for Children (WISC)* was compared by Tillman (1967a, b). The 100 students in each group ranged in age from 7 to almost 13 years and in IQ from 55 to 133. Performance was evaluated in terms of item-difficulty curves, t-tests of subtest means, and subtest reliabilities. Data were also analyzed to achieve a two-factor solution for each subtest. Results indicated that the blind scored about the same as the sighted on Arithmetic, Information, and Vocabulary; however, they did significantly less well on Comprehension and Similarities. Inclusion of Digit Span resulted in a higher mean IQ for the blind. Results of the factor analysis showed fewer factor loadings and weaker communalities for the blind on all subtests except Arithmetic, suggesting a greater specificity in the organization of abilities sampled by these tests. According to Tillman, the blind "tend to approach abstract conceptualization problems from a concrete and functional level and consequently lag behind the sighted children."

While verbal tests of varying quality have been available to measure mental ability, performance tests for blind children have been consistently

lacking. Dauterman, Shapiro, and Suinn (1967) reviewed performance tests of mental ability and discussed several problems related to their development.

Hecht and Newland (1965) examined the relationships between academic achievement and three measures of mental ability. Two of these were primarily verbal and one was a performance measure. While no single test appeared to be a better predictor of school achievement than any other, combination of either verbal measure with the performance measure resulted in better predictions. This study suffered from a woefully inadequate number of subjects.

Pointing out that many achievement tests contain items that are not suitable for braille, Nolan and Davis (1967) described how the most recent adaptations of the *Stanford Achievement Test (SAT)* and *Sequential Test of Educational Progress (STEP)* overcame this problem. Where items had to be omitted, the technique followed in both cases was to equate the distributions for the shortened tests with those for the sighted, norm population on the original tests. Trismen (1967) gave a technical description of this procedure for the *STEP* tests.

The vicissitudes of adapting tests for use by the blind were illustrated in two recent reports. Hammill and Powell (1967) described a relatively successful attempt to adapt the Abstraction Test for use with the blind. This test of cognitive ability uses auditory inputs exclusively and has previously been used with mentally retarded and cerebral palsied children. It appeared quite useful in discriminating levels of cognitive development in blind children between the ages of 5 and 10 years. Only one item proved unacceptable, and test reliability was high. In contrast, Gardner (1965) appeared to have suffered consistent frustration in his attempts to adapt the *Modern Language Aptitude Test* for use in predicting the success of blind students in a program to train Russian language translators.

Two recently developed tests dealt with braille reading. After preliminary research revealing the growth of ability to discriminate degrees of roughness, Nolan and Morris (1965) developed and validated the Roughness Discrimination Test as a measure of reading readiness for braille. The final form of the test (69 cards displaying squares of sandpaper of varying roughness) was found to have high reliability. Validities of the test for predicting reading speed and errors of 175 students at the end of grade 1 ranged from 0.53 to 0.57.

The construction of an individual Diagnostic Test of Braille Perceptual Skills was described by Hanley (1965). The purpose of this test was to measure perceptual errors peculiar to 215 letter, part-word, and whole-word symbols included in Grade 2 braille and to yield a profiled analysis of errors according to specific types of misperceptions and categories of braille orthography. Two equivalent forms of this test were developed, each containing the following five subtests; matching with recourse to a tactual stimulus, matching without recourse to a tactual stimulus, identification of



a symbol named, recognition of symbols, and the oral or written reproductions of symbols named. Reliabilities for both forms and all subtests were above 0.90. The tests discriminated between grade levels and were correlated positively with chronological age, number of years of braille usage, and intelligence.

### **Abstract Functioning and Concept Development**

Several research studies have compared the development of the cognitive abilities of the blind and the sighted. Juurmaa (1967) studied the structure of abilities as related to such factors as degree of remaining vision, age at onset of blindness, and duration of blindness. The abilities studied included verbal comprehension, tactual spatiality, arithmetic reasoning, musical aptitude, tactual discrimination, kinesthetic memory, and digital dexterity. The subjects used included a very heterogeneous group of 228 blind individuals and some undefined groups of sighted rehabilitation clients. Many subjects within the blind groups exceed the legal definition for blindness in the United States. With respect to degrees of blindness, there was no relation with verbal ability; a positive relation with numerical ability and memory (the severely blind were better); and a negative relation with spatial ability, tactual discrimination, and dexterity. With respect to spatial factors, it was demonstrated that the tests which formed a visualization factor with the sighted also determined a clear-cut factor on their own when they were performed tactually. Results of this study are difficult to evaluate as a whole because of the complexity, the possible lack of equivalence between visual and tactual forms of tests, and the varying numbers and kinds of subjects. However, as a pioneering effort, it possesses great heuristic potential.

The negative effects of visual handicaps on the development of spatial concepts were made more explicit in a dissertation by Hartlage (1967). He paired children blind from birth with sighted children for age, sex, and academic achievement. Included were 200 children from grades 2, 3, 5, 7, and 12. Subjects were required to respond to a test containing 16 items of a spatial nature and 16 items of a nonspatial nature. Eight questions of each type contained a self-referent and eight did not. Blind subjects' performance on spatial questions was inferior to that of the sighted Ss at all grade levels. Also, the blind performed less well on spatial questions as compared to nonspatial questions at all grade levels. The blind and the sighted performed equally well on nonspatial questions. All subjects performed better on questions containing a self-referent than on questions without a self-referent. The author concluded that vision plays an important role in the ability to deal with questions involving space.

The problem of verbalism in the blind has been one of enduring interest. Verbalism has been defined as excessive use of words that cannot be verified by concrete experience, or as the ability to define a word in an absence

of ability to identify by sensory means the object symbolized by the words. Docecki (1966) reviewed this problem in light of recent theorizing in the areas of language development and linguistics. He concluded that a concept of language restricted solely to word-thing relationships leads to an overstatement of the verbalism problem. He felt that there was no empirical evidence to lead one to assume that nonsensory-based language need disrupt the thinking process. While Docecki's conclusions are doubtlessly valid after cognitive development has reached a certain level, research on early stages of development reported during this review period raises serious questions as to its validity for young blind children.

Nagera and Colonna (1965) compared the ego and drive development of six blind children with that of sighted children. They concluded that whatever interfered with the child's ability to acquire adequate symbols could bring aspects of his ego development to a near standstill and result in extreme backwardness. In addition, they felt there was some evidence that lack of vision interfered with the ability to acquire and make use of words as symbols and that many essential elements of word symbols can be contributed only by sight. These psychoanalytically based conclusions are contrary to those of Docecki.

Research by Singer and Streiner (1966) also illuminated the verbalism area. They compared the richness and complexity of content of imaginative productions for matched groups of 20 blind and 20 sighted children. Sighted children were significantly higher for imaginativeness of play, spontaneous fantasy, and dreams. Blind children showed greater concreteness and lack of flexibility or associational variety. Fantasy themes were more perseverative. The authors concluded that the complexity and variety of stimulation available through sight was essential for fully-developed imaginative skill.

In a related study, Zweibelson and Barg (1967) required eight matched blind and sighted subjects to respond to words from the Similarities and Vocabulary subtests of the *WISC*. These responses were then categorized as concrete, functional, or abstract. Significant differences existed between the groups in that blind subjects did not use abstract concepts to the extent that sighted children did, but rather were found to think primarily on a concrete and functional conceptual level.

All of these studies point up the difficulties blind children have in reaching a cognitive level where abstract thinking is prevalent. Fixation at a concrete level of thinking implies that the number of associations attached to many verbal symbols will be restricted. This, of course, is reflected in their significantly lower scores on tests of abstract thinking (e.g., Tillman 1967a, b).

Regardless of the problems the blind encounter, word symbols remain the crucial raw material of the education process. The suggestion that undue emphasis on verbal skills can be debilitating to blind children was implied in a recent study by Kenmore (1965). Thirty matched pairs of

blind and sighted children from grades 3 and 6 were administered seven paired-associate learning tasks requiring use of different sense modalities. The tasks included verbal learning of meaningful and meaningless materials through reading and listening. Blind subjects handled familiar objects, random shapes, and tactually distorted objects while sighted subjects looked at color pictures of these items. For learning tasks involving familiar things, there were few differences among the groups regardless of the sense modality involved. For unfamiliar materials, however, the blind sixth graders excelled on the verbal materials but were inferior to all other groups in dealing with random shapes and distorted objects. Blind third graders exceeded blind sixth graders in tactual learning for these tasks.

Menaker (1966) investigated the effects of a prolonged deficit in the visual sense on development of perception in the tactile-kinesthetic sense. Tasks involving the kinesthetic aftereffect and the size-weight illusion were administered to matched groups of blind and sighted children who ranged in age from 5 to 15 years. Results showed a systematic change with age for the size-weight illusion, with the rate of change being slower for the blind.

## Education

### Listening

Interest in listening research remained high and appeared focused in these areas: (a) increasing learning efficiency through substituting listening for braille and large-type reading, (b) exploring uses of compressed speech as a means of furthering this aim, and (c) developing techniques for teaching listening skills.

Morris (1966) summarized the results of several studies which compared the relative efficiency of learning through listening with that of learning through reading braille or large type. Results of six studies involving both elementary and high school level subjects and materials from the areas of literature, social studies, and science were described. The major finding reported was that, at the elementary level, learning through listening and reading was equal. At the high school level, learning through reading exceeded that by listening for literature and social science, but not for science. However, when amounts learned were expressed in terms of times required to listen and to read, learning through listening appeared to be from 155 percent to 360 percent more efficient than that attained through reading either braille or large type. Comparative efficiency for listening was greater for braille reading than for large-type reading.

Research progress and research needs in the area of compressed speech were brought into sharp focus at the Louisville Conference on Time Compressed Speech (Foulke, 1967). Drawing participants from all over the United States, this conference dealt with technical problems in achieving speech compression and applications of speech compression in edu-



cation and other activities. Among the reports published in the proceedings of the conference was an excellent review of research by Foulke and Stitch (1967). Another report was that of de Hoop (1967), who studied the acquisition and retention of sentences presented at rates of 115, 210, 245, and 280 words per minute. The subjects included visually handicapped (VHC) and normally sighted children who were either of normal intelligence or mentally retarded. He found that, although performance of VHC and normally sighted children on initial learning and two-day retention was the same for sentences at 175 wpm, VHC excelled at rates of 210 wpm and 245 wpm. After two weeks, VHC excelled at retentions for 175 wpm, but the two groups were equal on retention for sentences at other rates.

Foulke (1966a) surveyed college students regarding the usefulness of compressed speech as a study medium. Samples of materials presented at rates of 180, 225, 275, and 350 wpm recorded on discs were sent to participants along with a questionnaire. Ninety-two percent of the respondents (50 percent of the participating group) indicated an interest in using compressed material. The most preferred rate was 275 wpm. Respondents felt compressed speech would be most useful for narrative and nontechnical exposition and least useful for novel or technical information. Most respondents found that they adapted readily to the use of compressed speech. In another study, Foulke (1966a) compared comprehension for materials compressed by the sampling method with that for those compressed through the speed changing method. A recording of a 3,350-word passage was reproduced at rates of 253, 300, and 350 wpm by both methods. Each of six groups of junior high level VHC listened to the passage in one form and then was tested for comprehension. Principal findings were that comprehension was inversely related to word rate with little comprehension occurring at 350 wpm and that there was no difference in comprehension for materials compressed by either of the two methods.

Two types of training designed to increase the listening comprehension of partially sighted students were compared in a doctoral dissertation by Bischoff (1967). Sixty-three partially sighted children were divided into three groups using random stratified procedures. All students were pre-tested with the *STEP Listening Test*, Form 3A, modified to eliminate the reading of the test items. Two groups received two 15-minute listening lessons per week for 10 weeks. Group one was trained using materials developed from *My Weekly Reader Listening Comprehension Paragraphs*. Group two was trained with materials from the *SRA Listening Lessons*. Group three received no training beyond the regular classroom schedule. All groups were post-tested with Form 3B of the *STEP Listening Test*. Both groups trained in listening demonstrated significant increases in listening comprehension scores following training. The control group showed no change. There was no difference in the degree of effectiveness of the two training methods.

## Tactile Reading

Consumer characteristics, perceptual processes, teaching techniques, code analyses, and raised line drawings have been the concern of recent research.

Goldish (1967) analyzed the market for braille. He suggested that there may be as many as 45,000 braille readers in the United States. Among these are 10,000 elementary and secondary school children and 1,500 college students. Nolan (1967), in a recent replication of past studies relating visual acuity and mode of reading, reported that, in 1966, 46 percent of legally blind children read braille. This is in contrast to 53 percent so reported in 1963. This decrease in braille readers appeared to be a consequence of a consistent trend in all school systems toward greater use of residual vision.

In a paper presented to the American Association of Instructors of the Blind, Nolan (1966) attempted to summarize the conclusions drawn from a six-year research project dealing with perceptual factors in braille word recognition. He concluded that unlike its counterpart in print reading, the process of braille word recognition appears to be a sequential integrative one in which word recognition is the result of the accumulation of pieces of information over a temporal interval. The perceptual unit is not the whole word shape, but the braille character. Skilled braille reading appears to be based on a probabilistic model of the braille reading environment. The complexity of the model developed is directly related to the mental ability of the reader. A number of perceptual cues are used to identify words correctly before all their individual characters are sensed. Among these are expectancies for the relative frequency of occurrence of letters in print, experience with the relative probabilities for letters following one another, experience with the occurrence of sets of letters, experience with the grammatical structure of the language, and cues stemming from preceding context. Subsequent research by others supported some of these conclusions.

If the braille character is the perceptual unit in word recognition, a logical inference is that training in character recognition should improve reading performance. Henderson (1966) tested this inference by giving 18 daily, individual sessions in character recognition to children in grades 3-6. During training, subjects were informed daily of the level of their performance, and remediation of errors was stressed. The results showed that the trained group reduced errors in character recognition by 85 percent and reduced mean character reading time by 30 percent. These changes far exceeded those found in a control group. Experimental subjects increased their oral reading rate by 12 wpm as compared to no increase for control subjects. In addition, it was found that 58 percent of the experimental as compared with 18 percent of the control Ss increased their rate of silent reading with no loss in comprehension. Kederis, Siems, and Haynes (1965) counted the frequency of occurrence of

braille symbols in print. They found a correlation coefficient of 0.46 between an index of ease of recognition for braille characters and their frequency of occurrence in print, substantiating the concept of a probabilistic model of reading environment devised by skilled braille readers.

Research by Pick, Thomas, and Pick (1966) confirmed the usefulness of experience with the relative probabilities for letters following one another as a cue in word recognition. They required braille readers to read pseudowords of from three to six letters. One-half the words were pronounceable and consisted of an initial consonant spelling, a vowel spelling, and a final consonant spelling. An unpronounceable counterpart was constructed for each pronounceable pseudoword. Findings were that unpronounceable words took significantly longer to read than pronounceable words. Significantly more errors were made in reading pronounceable words. Forty percent of errors involved addition or omission of one dot or letter, often making the pseudoword more pronounceable.

A study by Martin and Alonso (1967) pointed out the importance of cues from preceding context in braille reading. To determine whether braille material presented in telegraphic style could be learned as effectively as material presented in conventional style, these researchers required groups of subjects to read a traditional passage, a medium-length telegraphic version, and a short telegraphic version. Subjects reading the medium-length and short telegraphic versions required significantly less time to read than subjects reading the traditional version. However, the reading rate of subjects reading the short telegraphic version was significantly slower (40 percent) than that of those reading the traditional version. There were no differences among versions with respect to immediate or delayed recall of the materials.

Teaching braille reading is an area almost totalling unresearched. Lowenfeld and Abel (1967) surveyed programs for reading instruction in local and residential schools and studied the reading behavior of braille readers in grades 4 and 8. Replies to questionnaires by 88 percent of the schools solicited gave the following practices. Braille reading instruction generally starts in the first grade. Eighty-five percent of the teachers attempt to teach children to use both hands, with the remainder encouraging right hand use. Use of the index finger was emphasized. Braille writing is introduced with reading or shortly thereafter, and almost all teachers used the braillewriter.

One-third of the schools emphasized the braille alphabet in initial instruction while two-thirds emphasized whole-word reading and/or meaningful sentences. If the conclusions of Nolan (1966) are correct, this latter finding is of particular interest, since it implies that up to two-thirds of beginning braille instruction employs methods which violate the perceptual principles for word recognition. Nolan's conclusion would, of course, lead to the use of a decoding approach in initial reading instruction.



Lowenfeld and Abel used appropriate forms and subtests of the *STEP* and the *SAT* to measure reading performance of fourth- and eighth-grade subjects. Both reading speed and comprehension measures were obtained. General findings were that for reading comprehension, blind children were equal to seeing children in the fourth grade and superior at the eighth-grade level. The results for eighth graders are contrary to past findings. Since the average IQ of the eighth-grade group approximated 110, sampling problems may be a source of this ambiguity. Reading rate scores reported for both fourth-grade groups (local schools 84 wpm and residential schools 72 wpm) and eighth-grade groups (local schools 149 wpm and residential schools 116 wpm) greatly exceed those reported in any other studies of braille reading. Sampling may again be a problem; however, use of a one-minute reading speed test with no comprehension check may have also served to inflate scores.

Two studies evaluating the effectiveness of paced practice as a means of increasing braille reading speeds were reported by Kederis, Nolan, and Morris (1967). Subjects were braille readers from grades 5-12 in two residential schools. In one study, pacing consisted of requiring subjects to read familiar words, phrases, and short sentences as they were presented at diminishing time intervals by a controlled exposure device. In the second study, literary materials were presented at systematically increasing word rates through a moving belt display. No significant changes in reading speeds occurred as a result of practice. A serendipitous finding of great interest was that manipulation of motivational variables resulted in increases of reading speeds that averaged 25 percent and for some subjects reached 100 percent.

Not all tactile reading involves verbal symbology. Nonverbal symbology predominates in tactile graphs, charts, and maps. Schiff, Kaufer, and Mosak (1967) recently reported on several studies of tactual symbols for indicating directions. They compared traditional symbols which signaled direction (tactual drawings of visual arrows) with a specifically developed symbol which specified direction. The latter felt smooth when sensed in one direction, but sharp when felt in another. Results of their studies showed that while both symbols were equally effective in simple diagrams, use of the special symbol resulted in shorter response times and fewer errors as diagram complexity increased. Most subjects preferred the new symbol. In another study, Schiff and Isikow (1966) compared the readability of five-bar histograms which varied in combinations of difficulty of presentation and mode of presentation. Difficulty variables had significant effects on response time with medium differences in length giving the shortest times. Mode of presentation was not a significant variable in itself. However, as stimulus presentations became more difficult, the more redundant modes produced fewer errors. When stimulus presentations were easy, simple outline or textured modes were efficient.

## Sight Utilization

Recent studies in this area have concerned the use of large type by partially seeing children, the characteristics of large-type users, and the educational achievement of this group. In a national study involving more than 900 large-type users in grades 5 and 6, Birch and others (1966) attempted to achieve three goals: (a) to determine the effect of type size on the reading of the partially sighted, (b) to provide comprehensive information on school achievement based on reading the most favorable type size, and (c) to provide a better description of the characteristics of partially seeing children. All subjects were first given the five equivalent forms of a shortened version of the *Metropolitan Achievement Tests (MAT)* which had been reproduced in five different sizes of type. At the same time, personal data sheets for each participant were completed by their teachers. One year later, students who could be located were administered the full *Stanford Achievement Test* reproduced in the type size of that version of the *MAT* on which their raw score was highest.

The main contribution of the study lies in its description of a sizable group of large-type readers and possibly relevant variables. For example, it was found that large-type fifth- and sixth-grade pupils were overage for their grade by one year and nine months. In addition, these children were academically retarded by one year, giving an overall retardation of almost three years. With respect to far-point acuity, 28.7 percent fell above 20/70, 57.7 percent fell between 20/70 and 20/200, and 12.9 percent fell below 20/200. The study also revealed that teachers of many partially seeing children often lack data on the visual acuity and fields of their pupils and also on whether the students visual problems are stable or progressive. The authors combined their many findings into an interesting description of the typical partially seeing child.

A serious question can be raised as to the validity of the technique employed to identify the best type size for each pupil. About one-fifth of the group earned their highest *MAT* score on each of the five type sizes. This is the same distribution that would be expected on the basis of the simple random error which exists within five equivalent forms of the same test. There is no sure way of determining whether the obtained distribution is actually the result of type-size variation.

This study is an example of the possible influence of applications of computer technology on behavioral research. The authors adopted a brute-force approach to identification of relevant relationships by correlating almost every possible pair of variables. Since they did not describe the correlation coefficients used or the method for computing them, it is somewhat difficult to interpret the results. However, the majority of the coefficients reported approach zero and, consequently, the reader is relieved of this burden. The subsequent combination of 31 of the variables in a multiple regression equation to account for less than 17 percent of the

type-size variation is again a tribute to data processing technology but is of little practical import.

Bateman and Wetherell (1967) analyzed teacher responses to questionnaires covering 297 partially seeing children enrolled in grades 1-12 in order to identify educational characteristics of this group. They found that the amount of impairment was (a) directly related to dependence on large type; (b) inversely related to mental ability; and (c) unrelated to academic achievement, social-emotional adjustment, and parental attitude. The mental ability of the students was positively related to use of low-vision aids, acceptance into regular class, academic achievement, social-emotional adjustment, and positiveness of parental attitude.

## Mobility

A cursory review of research on mobility would lead to the conclusion that the previous emphasis on technological development had been continued; however, some evidence exists of a trend toward emphasis on training program development, training techniques, and evaluation.

Two recent studies represent examples of this trend. Leonard and Wycherley (1967) have attempted to establish generally acceptable criteria for travel performance and to cast these criteria in operational terms amenable to evaluation by a checklist. In the second phase of their work, the checklist was used to evaluate mobility proficiency for 59 subjects of varied backgrounds in mobility. The variables studied were success in navigating a specified route, rate of completion, ability to use landmarks, and ability to benefit from maps and to describe the route verbally. The data suggested that such skills are developmental in nature and that profiles of travel skills can be generated from the variables studied.

Lord (1967) has culminated a series of studies for the identification of orientation and mobility skills in the standardization of a scale for their appraisal in young blind children. The scale, considered an experimental edition, was based on revisions of 42 subscales now grouped into five major subscales: (a) self-help in relation to travel, (b) formal precane orientation and mobility skills, (c) movement in space, (d) use of sensory cues in travel, and (e) use of direction and making turns. Tentative norms were based on percents passing in specified age groups.

The studies by Leonard and Wycherley and by Lord are too broad in scope and too limited in populations, and they have methodological weaknesses. Nevertheless, they represent pioneering efforts critically needed for more objective evaluation of mobility techniques, training regimens, and program effectiveness.

Other significant aspects of mobility have been studied by Cratty (1966, 1967) and Cratty and others (1968). These studies have been focused on variables associated with the role of perception (i.e., perception of gradients, turning, and veering). It is encouraging to see the sustained interest on the part of an investigator reflected in Cratty's work.



## The Multiple Handicapped

Interest in the problems of multiple-handicapped blind children continued to expand rapidly during this review period (Ashcroft, 1966). Illustrative of local and national surveys of the problem are the works of Bucknam (1965), Wolf (1967), and Bennett and Oellerich (1966). Bucknam made an analysis of cumulative records, including medical examinations, for 137 children in a residential school for the blind, grades K-12 and ungraded. Seventy percent of the children had handicaps in addition to blindness, the most frequent being psychiatric disorders (49 percent). Thirty-three percent of the children were mentally retarded (i.e., possessing an IQ of less than 80) and 36 percent had brain injury. These data cannot be generalized to all schools for the blind, since such schools have widely differing admissions criteria. However, the data very probably represent a general increase in the incidence of such children and a trend toward increased willingness to admit these children to residential schools.

Wolf (1967) sought to determine how many mentally retarded blind children were enrolled in residential schools for the visually handicapped, their concomitant disabilities, admission criteria, and the extent and nature of educational services provided. Forty-eight heads of residential schools and 53 special class teachers in 28 of the schools responded to his extensive questionnaire. Unfortunately, local day schools, which enroll more than half of the visually handicapped children presently in school including large numbers of multiple-handicapped children, were not polled. Illustrative findings from the survey indicated that (a) children in residential school special classes for the mentally retarded had 3.2 presumptively diagnosed disabilities per child, one child reportedly having nine disabilities, (b) such classes were in reality classes for the multiply disabled rather than for a particular dyad of disabilities, and (c) the number of such children and programs for them are increasing. This study, utilizing one of the best and most comprehensive survey questionnaires, unfortunately had several limitations. Among these were its focus on mentally retarded blind children below 83 IQ, data which cannot be compared to previous surveys, and shifting frames of reference for reporting data on different aspects of the study.

The Bennett and Oellerich (1966) survey contacted 157 institutions for the mentally retarded and requested information on the number of blind or severely handicapped residents with mental retardation and the special programs or facilities for them. Most of the 96 reporting institutions enumerated 2 to 4 percent of the population as blind and mentally retarded, and 36 of the institutions had specially structured school programs for the visually handicapped.

Elonen, Polzien, and Zwarensteijn (1967) reported improvements in behavior, academic achievement, and measured intelligence for a small group of children (six) treated in a residential school for the blind. While

these results are encouraging, few generalizations are warranted because of the lack of a control group and the lack of specification of treatment and subjects.

Guess and Rutherford (1967) studied the effects of three types of stimulation on the stereotypical behavior (mannerisms or blindisms) frequently found in children with multiple handicaps. The experimental treatment reduced stereotyped behavior through a combination of objects (such as toys), sounds (such as recorded music and common sounds), and sound-generating devices (horns and bells). Sound alone failed to produce comparable results. This controlled study provided much-needed guidance regarding this aspect of behavior.

Several studies were designed to examine the effectiveness of methods useful with other groups. Templer and Hartlage (1965) tested the usefulness of the hand-face tests (i.e., subjects report body locus when touched by experimenters) with blind retardates and comparable, seeing retardates. The test failed to differentiate between the groups or to correlate significantly with mental age in the blind subjects. Hartlage (1965) found no significant difference in listening comprehension between a group of 12 retarded blind children matched with seeing peers. Cohen (1966) illustrated the value of a case study approach with a case history of a blind spastic child who, though also having been culturally and emotionally deprived, showed marked improvement with treatment. Seelye and Thomas (1966) explored the feasibility of providing mobility training for children with additional handicaps.

Though much remains to be done, the outlook is promising in terms of interest in the group, application of known techniques from other areas, and objective appraisal of the outcomes of treatment efforts.

## Summary and Conclusions

In their summary of the previous REVIEW, Ashcroft and Harley (1966) concluded, "A new era in research is needed to match the new era in education." It is not at all clear from the research reviewed during this period that the new era is upon us. Some of the research called for by these reviews has been undertaken. Assumptions basic to the braille reading process have been examined and brought into serious question. Intervention studies in the listening area have appeared. Concept development studies appear to be increasing. However, these represent only a beginning. Research in measurement still emphasizes adaptation. Human engineering approaches to development of aids and teaching materials have yet to become obvious. Applications of technological developments have yet to demonstrate practical results. To date there has been little evaluation of teacher and program effectiveness or teacher preparation. Consequently the "new era in research" remains a hope for the future.

## Bibliography

- AMERICAN FOUNDATION FOR THE BLIND. *Proceedings of the Rotterdam Mobility Research Conference*. New York: the Foundation (15 West 16th Street), 1965. 294 pp.
- ASHCROFT, SAMUEL C. "Delineating the Possible for the Multi-Handicapped Child with Visual Impairment." *International Journal for the Education of the Blind* 16: 52-55; December 1966.
- ASHCROFT, SAMUEL C., and HARLEY, RANDALL K. "The Visually Handicapped." *Review of Educational Research* 36: 75-88; February 1966.
- BATEMAN, BARBARA D. "Visually Handicapped Children." *Methods in Special Education*. (Edited by Norris Haring and Richard Schiefelbusch.) New York: McGraw-Hill Book Co., 1967. Chapter 9, pp. 257-301.
- BATEMAN, BARBARA D., and WETTERELL, JANIS L. "Some Educational Characteristics of Partially Seeing Children." *International Journal for the Education of the Blind* 17: 33-40; December 1967.
- BENNETT, FAY, and OELLERICH, D. W. "Institutional Facilities for the Visually Handicapped Mentally Retarded: Results of a Survey." *New Outlook for the Blind* 60: 233-35; October 1966.
- BIRCH, JACK W., and OTHERS. *School Achievement and Effect of Type Size on Reading in Visually Handicapped Children*. Cooperative Research Project No. 1766, Contract No. OEC-4-10-028. Pittsburgh: University of Pittsburgh, 1966. 166 pp. (Mimeo.)
- BISCHOFF, ROBERT W. *The Improvement of Listening Comprehension in Partially Sighted Students*. Doctor's thesis. Eugene: University of Oregon, 1967. 133 pp. (Abstract: *Dissertation Abstracts* 28: 2132A-33A; No. 6, 1967.)
- BUCKNAM, FRANK G. "Multiple-Handicapped Blind Children (An Incidence Survey)." *International Journal for the Education of the Blind* 15: 46-49; December 1965.
- COHEN, JEROME. "Development of a Blind Spastic Child: A Case Study." *Exceptional Children* 32: 291-94; January 1966.
- CRATTY, BRYANT J. "Perception of Inclined Plane While Walking Without Vision." *Perceptual and Motor Skills* 22: 547-56; April 1966.
- CRATTY, BRYANT J. "The Perception of Gradient and the Veering Tendency While Walking Without Vision." Research Bulletin No. 14. New York: American Foundation for the Blind (15 West 16th Street), 1967. pp. 31-51.
- CRATTY, BRYANT J., and OTHERS. "The Development of Perceptual-Motor Abilities in Blind Children and Adolescents." *New Outlook for the Blind* 62: 111-17; April 1968.
- DAUTERMAN, WILLIAM L.; SHAPIRO, BERNICE; and SUINN, RICHARD M. "Performance Tests of Intelligence for the Blind Reviewed." *International Journal for the Education of the Blind* 17: 8-16; October 1967.
- DE HOOP, WIETSE. "Effects and Interaction Effects of Speaking Role, Visual Limitation and Intelligence Level on Aural Acquisition and Retention of Sentences." *Proceedings of the Louisville Conference on Time Compressed Speech*. (Edited by Emerson Foulke.) Louisville, Ky.: University of Louisville, 1967. pp. 115-25.
- DOECKI, PAUL R. "Verbalism and the Blind: A Critical Review of the Concept and the Literature." *Exceptional Children* 32: 525-30; April 1966.
- ELONEN, ANNA S.; POLZIEN, MARGARET; and ZWARENSTEYN, SARAH. "The 'Uncommitted' Blind Child: Results of Intensive Training of Children Formerly Committed to Institutions for the Retarded." *Exceptional Children* 33: 301-307; January 1967.
- FOULKE, EMERSON. "Comparison of Comprehension of Two Forms of Compressed Speech." *Exceptional Children* 33: 169-73; November 1966. (a)
- FOULKE, EMERSON. "A Survey of the Acceptability of Rapid Speech." *New Outlook for the Blind* 60: 261-65; November 1966. (b)
- FOULKE, EMERSON, editor. *Proceedings of the Louisville Conference on Time Compressed Speech*. Louisville, Ky.: University of Louisville, 1967. 168 pp.
- FOULKE, EMERSON, and STITCH, THOMAS. "A Review of Research on Time Compressed Speech." *Proceedings of the Louisville Conference on Time Compressed Speech*. (Edited by Emerson Foulke.) Louisville, Ky.: University of Louisville, 1967. pp. 3-20.
- GARDNER, R. C. "A Language Aptitude Test for Blind Students." *Journal of Applied Psychology* 49: 135-41; April 1965.



- GOLDISH, LOUIS H. *Braille in the United States: Its Production, Distribution, and Use*. New York: American Foundation for the Blind (15 West 16th Street), 1967. 103 pp.
- GUESS, DOUG, and RUTHERFORD, GORIN. "Experimental Attempts to Reduce Stereotyping Among Blind Retardates." *American Journal of Mental Deficiency*. 71: 984-86; May 1967.
- HAMMILL, DONALD D., and POWELL, LaFAYETTE S. "An Abstraction Test for Visually Handicapped Children." *Exceptional Children* 33: 646-47; May 1967.
- HANLEY, LEO F. *The Construction of an Individual, Diagnostic Test of Braille Perceptual Skills*. Doctor's thesis. Boston: Boston University, 1965. 324 pp. Abstract: *Dissertation Abstracts* 28: 496A-97A; No. 2, 1967.
- HARTLAGE, LAWRENCE C. "Listening Comprehension in the Retarded Blind." *Perceptual and Motor Skills* 20: 763-64; June 1965.
- HARTLAGE, LAWRENCE C. *The Role of Vision in the Development of Spatial Ability*. Doctor's thesis. Louisville, Ky.: University of Louisville, 1967. 61 pp.
- HECHT, PATRICIA J., and NEWLAND, T. ERNEST. "Learning Potential and Learning Achievement of Educationally Blind Third-Eighth Graders in a Residential School." *International Journal for the Education of the Blind* 15: 33-38; December 1965.
- HENDERSON, FRED. "The Rate of Braille Character Recognition as a Function of the Reading Process." *American Association of Instructors of the Blind, 48th Biennial Conference*. Washington, D.C.: the Association, 1966. pp. 7-10.
- JOHNSON, DORIS J., and MYKLEBUST, HELMER R. *Learning Disabilities: Educational Principles and Practices*. New York: Grune and Stratton, 1967. 336 pp.
- JONES, JOHN W., and COLLINS, ANNE P. *Educational Programs for Visually Handicapped Children*. U.S. Department of Health, Education, and Welfare, Office of Education, Bulletin 1966, No. 6. Washington, D.C.: Government Printing Office, 1966. 74 pp.
- JONES, REGINALD L.; GOTTFRIED, NATHAN W.; and OWENS, ANGELA "The Social Distance of the Exceptional: A Study at the High School Level." *Exceptional Children* 32: 551-56; April 1966.
- JURMAA, J. *Ability Structure and Loss of Vision*. Research Series, No. 18. New York: American Foundation for the Blind (15 West 16th Street), 1967. 119 pp.
- KEDERIS, CLEVES J.; NOLAN, CARSON Y.; and MORRIS, JUNE E. "The Use of Controlled Exposure Devices to Increase Braille Reading Rates." *International Journal for the Education of the Blind* 16: 97-105; May 1967.
- KEDERIS, CLEVES J.; SIEMS, J. R.; and HAYNES, R. L. "A Frequency Count of the Symbology of English Braille Grade 2, American Usage." *International Journal for the Education of the Blind* 15: 38-46; December 1965.
- KENMORE, JEANNE R. *Associative Learning by Blind Versus Sighted Children with Words and Objects Differing in Meaningfulness and Identifiability Without Vision*. Doctor's thesis. Minneapolis: University of Minnesota, 1965. 169 pp. Abstract: *Dissertation Abstracts* 27: 959A; No. 4, 1966.
- LAWSON, L. J. "Ophthalmological Factors in Learning Disabilities." *Progress in Learning Disabilities*. Vol. 1. (Edited by Helmer R. Myklebust.) New York: Grune and Stratton, 1968. Chapter 7, pp. 147-81.
- LEONARD, J. ALFRED, and WYCHERLEY, R. J. "Towards the Measurement of Performance of Travel Skill." *Proceedings of the Conference for Mobility Trainers and Technologists*. Cambridge: Sensory Aids Evaluation and Development Center, Massachusetts Institute of Technology, 1967. pp. 27-45.
- LINVILL, J. G., and BLISS, J. C. "A Direct Translation Reading Aid for the Blind." *Proceedings of the IEEE*. Palo Alto, Calif.: Stanford University, 1966. pp. 40-51.
- LORD, FRANCIS E. *Preliminary Standardization of Scale of Orientation and Mobility Skills of Young Blind Children: Final Report*. U.S. Department of Health, Education, and Welfare, Office of Education, Bureau of Research, Project No. 6-2464, Grant No. OEG 4-7-062464-0369. Washington, D.C.: Government Printing Office, 1967. 148 pp.
- LOWENFELD, BERTHOLD, and ABEL, GEORGIE LEE. *Methods of Teaching Braille Reading: Final Report*. Project No. 2526, Contract No. OE-5-10-009. San Francisco: San Francisco State College, 1967. 93 pp. (Mimeo.)
- MARTIN, C. J., and ALONSO, L. *Comprehension of Full Length and Telegraphic Ma-*

- terials Among Blind Children: Final Report.* Educational Research Series, No. 42. East Lansing: Michigan State University, 1967. 52 pp.
- MENAKER, SHIRLEY L. *Perceptual Development in Blind Children.* Doctor's thesis. Boston: Boston University, 1966. 190 pp. Abstract: *Dissertation Abstracts* 27: 1625B-26B, No. 5, 1966.
- MORRIS, JUNE E. "Relative Efficiency of Reading and Listening for Braille and Large Type Readers." *American Association of Instructors of the Blind, 48th Biennial Conference.* Washington, D.C.: the Association, 1966. pp. 65-70.
- NACERA, HUMBERTO, and COLONNA, ALICE B. "Aspects of the Contribution of Sight to Ego and Drive Development; A Comparison of the Development of Some Blind and Sighted Children." *The Psychoanalytic Study of the Child*. Vol. 20. (Edited by Ruth Eissler and others.) New York: International Universities Press, 1965. pp. 267-87.
- NATIONAL ADVISORY COMMITTEE ON HANDICAPPED CHILDREN. *Special Education for Handicapped Children.* U.S. Department of Health, Education, and Welfare, Office of Education. Washington, D.C.: Government Printing Office, 1968. 44 pp.
- NOLAN, CARSON Y. "Perceptual Factors in Braille Word Recognition." *American Association of Instructors of the Blind, 48th Biennial Conference.* Washington, D.C.: the Association, 1966. pp. 10-14.
- NOLAN, CARSON Y. "A 1966 Reappraisal of the Relationship Between Visual Acuity and Mode of Reading for Blind Children." *New Outlook for the Blind* 61: 255-61; October 1967.
- NOLAN, CARSON Y., and DAVIS, C. J. "Measurement of Scholastic Attainment with Standardized Tests." *Proceedings of the Fourth Quinquennial Conference of the International Council of Educators of Blind Youth.* Watertown, Mass.: Perkins School for the Blind, 1967. pp. 86-92.
- NOLAN, CARSON Y., and MORRIS, JUNE E. "Development and Validation of the Roughness Discrimination Test." *International Journal for the Education of the Blind* 15: 1-6; October 1965.
- PICK, ANN D.; THOMAS, MARGARET L.; and PICK, HERBERT L., JR. "The Role of Grapheme-Phoneme Correspondences in the Perception of Braille." *Journal of Verbal Learning and Behavior* 5: 298-300; June 1966.
- SCHIFF, WILLIAM, and ISKOW, HERBERT. "Stimulus Redundancy in the Tactile Perception of Histograms." *International Journal for the Education of the Blind* 16: 1-11; October 1966.
- SCHIFF, WILLIAM; KAUFER, L.; and MOSAK, S. "Informative Tactile Stimuli in the Perception of Direction." Research Bulletin No. 14. New York: American Foundation for the Blind (15 West 16th Street), 1967. pp. 65-94.
- SEELYE, WILMA S. and THOMAS, JOHN E. "Is Mobility Feasible with Multiply Handicapped Blind Children?" *Exceptional Children* 32: 613-17; May 1966.
- SENSORY AIDS EVALUATION and DEVELOPMENT CENTER. *Proceedings—Conference on New Processes for Braille Manufacture.* Cambridge: Massachusetts Institute of Technology, 1969. 76 pp.
- SINGER, JEROME L., and STREINER, BELLA F. "Imaginative Content in the Dreams and Fantasy Play of Blind and Sighted Children." *Perceptual and Motor Skills* 22: 475-82; April 1966.
- TEMPLER, DONALD, and HARTLACE, LAWRENCE. "The Reliability and Utilization of the Hand-Face Test with the Retarded Blind." *American Journal of Mental Deficiency* 70: 139-41; July 1965.
- TILLMAN, M. H. "The Performance of Blind and Sighted Children on the Wechsler Intelligence Scale for Children: Study I." *International Journal for the Education of the Blind* 16: 65-74; March 1967. (a)
- TILLMAN, M. H. "The Performances of Blind and Sighted Children on the Wechsler Intelligence Scale for Children: Study II." *International Journal for the Education of the Blind* 16: 106-12; May 1967. (b)
- TRISMEN, DONALD A. "Equating Braille Forms of the Sequential Tests of Educational Progress." *Exceptional Children* 33: 419-24; February 1967.
- WOLF, J. M. *The Blind Child with Concomitant Disabilities.* Research Series, No. 16. New York: American Foundation for the Blind (15 West 16th Street) 1967. 91 pp.
- ZWEIBELSON, I., and BARG, C. FISHER. "Concept Development of Blind Children." *New Outlook for the Blind* 61: 218-22; September 1967.

## CHAPTER V

### Learning Disabilities

CORRINE E. KASS\*

Learning disabilities has gained wide-spread acceptance as a professional label despite the continuing problem of providing an adequate definition of the term. In an effort to clarify the concept of learning disabilities for the special educator, an institute for advanced study was funded by the Bureau of Education for the Handicapped, U.S. Office of Education, and held at the Institute for Language Disorders, Northwestern University, Evanston, Illinois, in 1967. The resulting definition, written by the fifteen invited participants (all from special education), goes beyond the description of characteristics of children with learning disabilities and suggests that there are educational methodologies which will change discrepancies in functioning. When a special methodology is applied following appropriate diagnosis, the assumption is made that the child will learn in an essentially normal manner. The resultant definition includes the following:

- Learning disability refers to one or more significant deficits in essential learning processes requiring special education techniques for remediation.
- Children with learning disability generally demonstrate a discrepancy between expected and actual achievement in one or more areas, such as spoken, read, or written language, mathematics, and spatial orientation.
- The learning disability referred to is not primarily the result of sensory, motor, intellectual, or emotional handicap or lack of opportunity to learn.
- *Significant deficits* are defined in terms of accepted diagnostic procedures in education and psychology.
- *Essential learning processes* are those currently referred to in behavioral science as involving perception, integration, and expression, either verbal or nonverbal.
- *Special education techniques for remediation* refers to educational planning based on the diagnostic procedures and results.

### Research in Transition

For many years following the initial work of Alfred Strauss, research questions focused on characteristics of the "Strauss syndrome" child, such as (a) perceptual disorders, (b) perseveration, (c) thinking or conceptual disorders, (d) behavioral disorders, (e) slight neurological signs, (f) his-

\* The author wishes to acknowledge the valuable help of John Dewitt, graduate assistant, University of Arizona, in the preparation of this manuscript.



tory of neurological impairment, and (g) no history of mental retardation in the family. Numerous hypotheses were suggested regarding the brain functions or dysfunctions of such children. The range of these studies was from neurology to educational methodology. The methods devised by Strauss and his co-workers are still used in classes for brain-injured children. Recently, however, research questions of special educators regarding this particular set of symptoms seem to be waning, although interest in the learning problems of a given population has been mounting. Public school services for "Strauss syndrome" children have never really been instituted as widely as needed.

In their efforts to share the responsibility for understanding and serving children with special learning disabilities, the several disciplines have had difficulty deciding on research questions relevant to their respective roles. The literature reveals a large number of so-called "white papers," reflecting a transition toward asking different research questions based on different theoretical formulations. This trend is reflected in the small number of research reports on learning disabilities found in the journals. *Exceptional Children* and the new *Journal of Learning Disabilities* contain many articles relating to theory, not to research. *Perceptual and Motor Skills* and *The Slow Learning Child* are exceptions in that more research appears in these than in other journals.

The few books of readings which have appeared since the last REVIEW (Bateman, 1966) are compilations showing substantially more position papers than research reports (Cruickshank, 1966; Frierson and Barbe, 1967; Hellmuth, 1966; Hellmuth, 1968; Money, 1966; and Myklebust, 1967). The one exception is a complete textbook by Johnson and Myklebust (1967) in which their particular theory of learning disabilities has led to a number of research questions, including the question of prevalence. These are presently being studied in several multidisciplinary projects. All the results have not yet been reported in complete form.

The research chosen for review in this issue reflects the desire of the educator to follow different lines of questioning. The studies to be reviewed have been selected because they attempt to answer three major research questions: (a) What are some correlates of learning disabilities? (b) What special education methods will ameliorate or prevent learning disabilities? (c) What research methodology is required in order to relate effective remedial procedures to correlates of learning disabilities?

## Correlates of Learning Disabilities

### Prediction Correlates

A predictive index was developed by de Hirsch, Jansky, and Langford (1966) for the purpose of identifying children at the kindergarten level who presented a specific pattern of dysfunction found to be related to

developmental lag. The following tests comprised the index: (a) Pencil Use, (b) *Bender Visual Motor Gestalt Test for Children*, (c) *Auditory Discrimination Test*, (d) Number of Words Used in a Story, (e) Categories, (f) *Reversal Test*, (g) Word Matching Test, (h) Word Recognition I, (i) Word Recognition II, and (j) Word Reproduction.

The Dyslexia Schedule was devised by McLeod (1966) for the purpose of assessing aptitude for childhood dyslexia. This predictive device was used as the nucleus of an interview between a social worker and the parents. This validation study was done in Brisbane, Australia, where the incidence of reading retardation was found to be low (3.25 percent as compared to 21 percent in Britain). As the author points out, we might expect that incidence of dyslexia in the neurological or genetic sense would be relatively constant from culture to culture. Therefore, the child with reading retardation in Brisbane would be more likely to be dyslexic than would one in a place where the incidence of reading retardation is greater. McLeod found a significant difference between the experimental and the control groups in 23 of the 90 original items. A large-scale investigation with this instrument is being planned to determine whether it can be used as a screening device.

Sapir and Wilson (1967) measured developmental performance of kindergarten-age children in an effort to identify problems in perceptual-motor skills, bodily schema, and language development. Out of 64 children in this pilot study, the Sapir Developmental Scale classified 18 children as developmentally deficient in at least two of the three areas. A comparison with neurological diagnosis of minimal cerebral dysfunction showed significant correlation. Further validation of the developmental scale with later academic performance also showed significant correlations. Trained kindergarten teachers can administer this test which takes about 30 minutes.

Tauber (1966-67) did an identification study with 30 kindergarten children and found three tests significantly predictive of difficulty in mastering fundamental skills: (a) *Knox Cube Test*, (b) *Verbal Language Development Scale*, and (c) *Auditory Discrimination Test* (dissimilar word-pair items). The last two tests showed a combined multiple correlation of 0.92, and Tauber suggested that these two tests should be studied on a larger school population.

Carró (1968) obtained preliminary findings on an instrument designed to assess the development of auditory comprehension of language structure (lexical and grammatical). On this test, the subject pointed to the picture in each item corresponding to the linguistic utterance of the examiner. This test is similar to the Auditory-Vocal Automatic subtest (grammar) of the *Illinois Test of Psycholinguistic Abilities (ITPA)*. Mean language comprehension scores increased with age most rapidly between ages of 3 and 4½ years. According to this analysis, one category of form classes or function words does not develop before another; rather, auditory comprehension of language structure depends on the particular linguistic structure, its refer-

ent, and frequency of use. Haring and Ridgway (1967) found that the screening observations of kindergarten teachers were as predictive of learning disabilities as any means of assessment the investigators had selected.

### Age-Related Correlates

Belmont and Birch (1966), in a carefully done epidemiologic study, analyzed the intellectual patterning on the *Wechsler Intelligence Scale for Children (WISC)* for 150 retarded readers and 50 normal readers. In an attempt to correct methodological weaknesses of previous studies which relate *WISC* patterning to reading performance, Belmont and Birch (a) selected their sample from an entire community not a clinical population, (b) selected a population homogeneous in age and sex (9-year-old boys), and (c) selected a comparison group of normal readers from the same school classes, from the same distribution of social class, and from the same birth year as the sample population. Retarded readers were identified as having scores below the 10th percentile on three or four out of four reading tests. This resulted in a mean difference between the groups of 1.7 school years.

The *WISC* patterns were successively analyzed with the following results: (a) the entire group of retarded readers had significantly lower IQ's than the normal readers, (b) when borderline intellectual level was eliminated (IQ: 80-89), the retarded readers were significantly lower in verbal IQ than the normal readers, (c) when the two groups were matched for average intelligence (IQ: 90-109), the retarded readers were again significantly lower on IQ, (d) poorest readers from the retarded readers group (below 10th percentile on all four reading tests) were significantly higher than normal readers on performance IQ and significantly lower on verbal IQ, and (e) analysis of items in the vocabulary subtests suggested that a correlate of reading disability at age 9 is inadequacy in language functioning rather than in perception.

It must be pointed out that these findings can be generalized only to children of the same age and grade level (grade 4). While the mastery of perceptual skills may be important in the early stages of reading development, the conceptual factors noted in this study may be more important for later reading.

Studies of the relationship of auditory and visual pattern recognition to reading were conducted by Sterritt and Rudnick (1966) on fourth-grade boys and by Rudnick, Sterritt, and Flax (1967) on third-grade boys. At the fourth-grade level, the only modality related to reading was the auditory, while at the third-grade level, both auditory and visual modalities were found to be related to reading, suggesting that, with increasing age, auditory functions in language become increasingly important in reading. The verbal mediation necessary for performance on the *WISC* verbal section may also be necessary in auditory pattern recognition.



Reed (1967a) located 38 children in a public school first grade and 38 children in a fifth grade who showed the *WISC* profile of high performance IQ and low verbal IQ. He also found 16 children in grade 1 and 25 children in grade 5 who had higher verbal IQ's than performance IQ's. A third group of children were those with equal verbal and performance IQ's. Reed confirmed the findings of Belmont and Birch (1966) that high verbal—low performance children are good readers, but he did not find, as did Birch and Belmont (1964), that high performance—low verbal IQ's were significantly related to poor reading. The population was not selected because of learning difficulties, and in this population depressed verbal ability did not result in poor reading. Reed questioned the feasibility of predicting poor reading from intellectual patterns and suggested that fluent reading may itself lead to superior verbal skills.

Reed (1968) identified good and poor readers in grades 1 and 5 on the basis of reading scores. His question was, "Does knowledge of the ability deficits of older children who have difficulty in reading help us to prevent younger children from becoming reading failures?" On the basis of the *WISC* and other tests, he concluded that older children appeared to have more difficulty with the symbolic significance of language while younger children have more difficulty in visuo-spatial relations. Therefore, it appears that knowledge of the language disabilities of older children is not particularly useful for the development of the skills in early reading.

The relationship between "word blindness" (dyslexia) and certain factors, such as finger localization and laterality, has been confirmed in certain studies and not in others. That age-related factors also have relevance is indicated in a study by Reed (1967b), in which he tested the finger localization of all children enrolled in grades 1 and 5 in three community schools. Difficulty in finger localization is the inability to differentiate among the fingers on the basis of tactile cues alone. No special implications can be drawn from the age-6 sample, but at age 10 there were significant differences in right-hand errors between poor readers and good readers. The results indicate that while the 10-year-old children with reading deficits did not have clinically demonstrable brain lesions we can infer that they had some sort of disturbance in the integrity of the left cerebral hemisphere.

Silver and Hagin (1966) reported a follow-up study of 18 children with specific reading disabilities seen at a hospital clinic. These children had been diagnosed initially between 1949 and 1951 and retested periodically. When the subjects were between ages 16 and 24, an analysis of retests of perceptual abilities and reading achievement was made. Results were reported for two subgroups: (a) the organic group (classical neurological deficits) showed few changes in perceptual abilities over time, particularly in accuracy of visual perception and figure background, and (b) the developmental group (basic defect in capacity to integrate written material and to associate concepts with symbols) showed striking improvement in

perceptual abilities, except for continuing laterality difficulties. The organic group displayed poorer reading comprehension at the adult level than the developmental group did.

Three implications can be drawn from the project results: (a) perception deficits change over time, (b) perceptual maturation must be considered in long-range planning, and (c) training methods for children with organic problems must be improved.

## Psychological Correlates

The correlation which Birch and Belmont (1964) found between inter-sensory integration deficits and poor reading was the stimulus for the following two studies:

Beery (1967) used Birch and Belmont's test of Auditory-Visual Integration in a study of 15 children of normal intelligence with specific reading disability and an equal number of controls matched for IQ, sex, and age. She devised an additional test of Visual-Auditory Integration in which she reversed the stimulus and response modalities. The visual stimulus (dots on a card) preceded the auditory stimulus (tones matching the dot patterns). It was suggested that Visual-Auditory Integration might be more closely related to reading because of the visual nature of reading. What Beery found, however, was that the performance of dyslexic children was inferior to that of the controls on both tests. The age range in this study was relatively wide—8 to 13 years.

Blank and Bridger (1966) studied the Birch and Belmont conclusion that poor readers' difficulty with cross-modal transfer may be related to their inability to apply verbal mediators to the auditory and visual stimuli. Thirteen poor readers and thirteen good readers were matched for intelligence and age (9 years), and were then compared for their ability to convert auditory taps to visual dot patterns. The retarded readers were significantly poorer in the ability to convert temporal stimuli (flashing-light patterns) to spatial stimuli (dot patterns on paper) and in the accuracy with which they were able to apply verbal labels to the light flashes. These researchers agreed with Birch and Belmont that a general deficit in children with reading disability may be the inability to apply correct verbal labels to temporally presented stimuli.

McLeod (1968) noted integrative and memory correlates when he administered the *WISC*, *ITPA*, and some auditory and written tests to 23 7-year-old children who were more than one year retarded in reading relative to CA. A control group was also tested. A factor analysis of the data revealed four significant factors in reading retardation.

In a federally funded research project, Rubin and Braun (1967) reported the use of a behavior checklist on which teachers rated 400 children demonstrating problem behaviors from grades 1, 2, 3, and 5. A factor analysis showed that the factor Disoriented Behavior was related to cognitive-motor

dysfunction (poor coordination and poor perceptual awareness) for some children with behavior problems but not for others. It was proposed that the first group had secondary emotional disturbance while the second group had primary emotional disturbance. Prediction on the basis of this instrument will lead to better understanding of one of the basic differences between emotional disturbance and learning disabilities.

### Neurochemical Correlates

Since the initial shock of the Smith and Carrigan report (1959), very little publicity has been given to the chemical imbalance factors suggested by Smith. One reported study by Wilhelm (1966) took some of the case studies from the original study of children with severe reading disability. These cases fit within the "muscular tension" syndrome of the original study. This syndrome involved such signs as unusual need for milk, suppressed irritability, and both auditory and visual deficits. When the channel memory deficits were compared on the *Knox Cube Test* and the Digit Span subtest of the *WISC*, these children were found to have a lower score on the Digit Span of at least 0.7 standard score. Smith suggested that "constant input of nonspecific or random stimulation to the cortex would result from muscle irritability and thus might interfere with the storing process thought to account for memory." This study is interesting in that auditory memory is the primary diagnostic sign. Implications for remediation suggest that the child displaying this type of disability ought to have training in listening as well as in visual discrimination. The researcher concluded that this syndrome may be part of a more general communication problem.

The question of biochemical influence on learning is one which might have a great deal of value for education once treatment is developed for children with chemical irregularities. Myklebust (1967) is in the process of determining some of these relationships at the neurochemical level. Developmentally, learning disability research is in a transitional stage. While specific test patterns are appearing and certain neurochemical aberrations are being confirmed, remedial treatments cannot yet be directly traced to the diagnostic findings as reported in the above studies.

### Conclusions

The question of correlates of learning disabilities is best studied specifically, that is relative to prediction, age, behavior, and neurochemistry.

Prediction of learning disabilities was initially made at the kindergarten level in the studies which were reviewed. The predictive procedures were significant within a range from an extensive battery of standardized tests to teacher evaluation. While teacher prediction seemed effective, an unanswered question is whether there is a correlation between prediction effectiveness and prevention. Predicting later achievement difficulties from



given characteristics at around the five-year level may, at best, be something which any relatively well trained observer, with or without measuring instruments, can do and, at worst, something which merely attaches an "expectation" label in school.

*Age-related correlates* of learning disabilities have acquired more significance since Belmont and Birch (1966) controlled the age factor in their epidemiologic study. Characteristics which are ascribed to learning disabilities at one age (for example, as noted in the prediction studies) are not always the same as at another age.

The Spearman-type question of general versus specific factors is relevant and must be considered by the researcher. While certain characteristics appear in descriptions of children with learning disabilities again and again, whether these are discrete factors is still open to question.

Quereshi (1967) tested changes in the organization of cognitive psycholinguistic development over the span of  $2\frac{1}{2}$  through 9 years of age. The sample was comprised of 700 children of the standardization sample of the *ITPA*. A factor analysis of the *ITPA* revealed a general ability factor and some group factors. Quereshi found that group factors assume importance only with increasing age and recommended that tests intended for children up to 10 years of age should be so constructed as to emphasize the general factor during the preschool years and only gradually give greater prominence to group factors. Such a conclusion suggests that relatively global measures of ability are adequate identification or prediction measures for preschool children.

*Psychological and neurochemical* correlates of learning disabilities were reported in relatively few studies over the past three years and few correlates were noted. These were given such labels as "intersensory integration," "verbal mediation," "sequencing," "disoriented behavior," and "channel memory."

## Special Education Methods

The studies under this heading are isolated attempts to validate specific methods which will ameliorate or prevent learning disabilities and which must be analyzed within a different context from the previous studies.

Painter (1966) evaluated 20 children with low-average IQ's in sensory, motor, and spatial performance skills and gave 21 training sessions to half the group. The training session methods brought about significant gains in perceptual-motor integration for the experimental group, showing that a remediation program geared to specific deficits does result in improvement in those abilities. The relationship to later school learning was not studied.

Gardner (1966) reported several field trials of the use of i.t.a. with eight 10-year-old pupils who had failed to learn to read after a minimum

of two years in school. After eight weeks of four sessions with a special i.t.a. teacher each week, the experimental pupils were superior to the control groups on a traditional reading test of word recognition. After 30 weeks, it was noted that the gains continued and that the acquired skills transferred completely to traditional reading. For some pupils, however, it was necessary to give auditory and visual training before i.t.a. could be learned. It was also found that remedial teaching took less time before discharge with i.t.a. than with the traditional method (14 months on the average as compared to 18 months). When the groups were compared two years later, it was noted that 85 percent of the i.t.a. pupils continued to make satisfactory progress while 70 percent of the control pupils had made satisfactory progress, with the exception of pupils with IQ's of less than 85.

Jacobs (1968) evaluated the Frostig Visual Perception Training Program with prekindergarten, kindergarten, and first-grade children. Only the experimental group of first graders did significantly better on the Frostig post-test than their controls. It may be that the Frostig Program adds something to the first-grade curriculum, but is already in the curriculum for prekindergarten and kindergarten. The population included a large proportion of disadvantaged children.

Silver, Hagin, and Hersh (1967) reported one portion of their detailed analysis of 80 boys referred to a hospital clinic for school-learning and behavior problems. This was a cross-over experiment using two methods of remediation in alternating six-months blocks of time. The two methods were (a) the method of stimulation of deficit perceptual areas, designed to develop accuracy of perception within a given modality and (b) contact appointments, involving individual work with school books and tutoring in school subjects. One group received the first method for the first six months, and the second method for the second six months of the experiment. Group two received the second method for the first six months and the first method for the second six months of the experiment. A representative case study from each of the two groups reflected the results—more progress in achievement test scores was noted following the perceptual training phase (the first method) especially when this was provided first. The perceptual training methods included 41 techniques which are described in an instructional manual.

Wilson (1968) investigated three clinical techniques (traditional method, linguistic approach, and experience approach) applied to children with associative learning difficulty stemming from three possible causes (educational, psychological, and organic). The three hypotheses related the three clinical techniques to the three causes: (a) the traditional basal reader approach would be most successful where the cause of the learning difficulty was educational; (b) the linguistic approach would be most successful where the cause was psychological in origin; and (c) the experience approach, exemplified by the Fernald method, would be most successful where the cause was organic. The sample was too small, the teaching time too

limited, and the statistical procedure too imprecise to comment on the research as such. It is reported here because the hypotheses are considered significant for further research.

## Conclusions

More studies could have been included in this section, but research projects in teaching methods usually show a frustrating lack of controlled variables. The methods which appear to be most effective are prescriptive in nature, that is, they are specific techniques applied to specific deficits. Improvement in specific areas of perception cannot be generalized to all of perceptual functioning. The complexity of perceptual functioning is understandable when the probabilities of specific deficits appearing in combination in any one child are considered. No matter how much special educators decry the panaceas which appeal to parents of children with learning disabilities, it must be admitted that there is not yet enough research information for accepting or rejecting most of these methods.

## Research Methodology

Traditional experimental designs have not always been useful for bridging the gap between diagnosis and remediation. Kirk (1966) suggested the use of the case study for the purpose of first deriving generalizations from a single case and then confirming them by further experimentation. A single study by itself does not constitute grounds for generalization, except as it is confirmed through a number of case studies. In the Kirk report, nine children were given the *ITPA* and other tests and then were tutored and retested. The tentative generalizations from the case studies were that (a) some children have test scores within the mentally retarded range because of discrepancies in psycholinguistic functioning, (b) remediation is effective when related to specific psycholinguistic deficits, (c) remediation of deficits is more effective when applied to preschool children than to school-age children, and (d) appropriate remediation varies from one case to another.

Another form of case-study research is the analysis of genetic traits over a number of generations. McGlannan (1968) reported a pilot study of the genetic traits of 65 families each having one or more children with two-year deficiencies in language and/or reading skills. The sample population did not have demonstrable organicity, psychopathic or neurotic evidence, or uncorrected sensory modality defects. From an intensive selection process, which involved specific test patterns and membership in a family unit (defined as three family generations) where two or more persons presented language disorders, McGlannan noted (a) a vulnerable-family syndrome characterized by ambilaterality, directional confusion, and maturational lags, and (b) underdevelopment at birth and vulnerability to



results of perinatal trauma. McGlannan suggested that prevention of certain learning disabilities may be effected through education of vulnerable families. This pilot study will be followed by a longitudinal familial investigation called Project Genotype.

## Conclusions

It is through the imaginative restructuring of experimental designs that special educators may expect to achieve relevant remedial procedures. Contrasted with the curriculum in the area of mental retardation, remediation in learning disabilities is based on the assumption that the child could function in the regular classroom given appropriate special education.

## Bibliography

- BATEMAN, BARBARA. "Learning Disorders." *Review of Educational Research* 36: 93-119; February 1966.
- BEERY, JUDITH WILLIAMS. "Matching of Auditory and Visual Stimuli by Average and Retarded Readers." *Child Development* 38: 827-33; September 1967.
- BELMONT, LILLIAN, and BIRCH, HERBERT G. "The Intellectual Profile of Retarded Readers." *Perceptual and Motor Skills* 22: 787-816; June 1966.
- BIRCH, HERBERT G., and BELMONT, LILLIAN. "Auditory-Visual Integration in Normal and Retarded Readers." *American Journal of Orthopsychiatry* 34: 852-61; October 1964.
- BLANK, MARION, and BRIDGER, WAGNER. "Deficiencies in Verbal Labeling in Retarded Readers." *American Journal of Orthopsychiatry* 36: 840-47; October 1966.
- CARROW, SISTER MARY ARTHUR. "The Development of Auditory Comprehension of Language Structure in Children." *Journal of Speech and Hearing Disorders* 33: 99-111; May 1968.
- CRUICKSHANK, WILLIAM M., editor. *The Teacher of Brain-Injured Children*. Syracuse, N.Y.: Syracuse University Press, 1966. 334 pp.
- DE HIRSCH, KATRINA; JANSKY, JEANNETTE J.; and LANGFORD, WILLIAM. *Predicting Reading Failure*. New York: Harper and Row, 1966. 144 pp.
- FRIERSON, EDWARD, and BARBE, WALTER, editors. *Educating Children with Learning Disabilities: Selected Readings*. New York: Appleton-Century-Crofts, 1967. 502 pp.
- GARDNER, KEITH. "The Initial Teaching Alphabet (i.t.a.) and Remedial Reading Programme." *The Slow Learning Child* 13: 67-71; November 1966.
- HARING, NORRIS G., and RIDGWAY, ROBERT W. "Early Identification of Children with Learning Disabilities." *Exceptional Children* 33: 387-95; February 1967.
- HELLMUTH, JEROME, editor. *Learning Disorders*. Vol. II. Seattle, Wash.: Special Child Publications, 1966. 423 pp.
- HELLMUTH, JEROME, editor. *Learning Disorders*. Vol. III. Seattle, Wash.: Special Child Publications, 1968. 605 pp.
- JACOBS, JAMES N. "An Evaluation of the Frostig Visual-Perceptual Training Program." *Educational Leadership* 1: 332-40; January 1968. (Research Supplement)
- JOHNSON, DORIS J., and MYKLEBUST, HELMER. *Learning Disabilities*. New York: Grune and Stratton, 1967. 336 pp.
- KIRK, SAMUEL A. *The Diagnosis and Remediation of Psycholinguistic Disabilities*. Urbana: Institute for Research on Exceptional Children, University of Illinois, 1966. 250 pp.
- MCGLENNAN, FRANCES K. "Familial Characteristics of Genetic Dyslexia: Preliminary Report from a Pilot Study." *Journal of Learning Disabilities* 1: 185-91; March 1968.
- MCLEOD, JOHN. "Prediction of Childhood Dyslexia." *The Slow Learning Child* 12: 143-55; March 1966.

- MCLEOD, JOHN. "Reading Expectancy from Disabled Learners." *Journal of Learning Disabilities* 1: 97-105; February 1968.
- MONEY, JOHN, editor. *The Disabled Reader*. Baltimore, Md.: Johns Hopkins Press, 1966. 421 pp.
- MYKLEBUST, HELMER, editor. *Progress in Learning Disabilities*. Vol. I. New York: Grune and Stratton, 1967. 273 pp.
- NORTHWESTERN UNIVERSITY and U.S. OFFICE OF EDUCATION. "Proceedings of the Conference on Learning Disabilities and Interrelated Handicaps." Evanston, Ill.: Northwestern University, 1967. 267 pp. (Unpublished) (Mimeo.)
- PAINTER, GENEVIEVE. "The Effect of a Rhythmic and Sensory Motor Activity Program on Perceptual Motor Spatial Abilities of Kindergarten Children." *Exceptional Children* 33: 113-16; October 1966.
- QUERESHI, MOHAMMED Y. "Patterns of Psycholinguistic Development During Early and Middle Childhood." *Educational and Psychological Measurement* 27: 353-65; Summer 1967.
- REED, JAMES C. "Lateralized Finger Agnosia and Reading Achievement at Ages 6 and 10." *Child Development* 38: 213-20; March 1967. (a)
- REED, JAMES C. "Reading Achievement as Related to Differences Between *WISC* Verbal and Performance IQ's." *Child Development* 38: 835-40; September 1967. (b)
- REED, JAMES C. "The Ability Deficits of Good and Poor Readers." *Journal of Learning Disabilities* 1: 134-39; February 1968.
- RUBIN, ELI Z., and BRAUN, JEAN S. *Behavioral and Learning Disabilities Associated with Cognitive-Motor Dysfunction*. U.S. Office of Education, Bureau of Research, Interim Report. Detroit, Mich.: Lafayette Clinic, 1967. 8 pp. (Mimeo.)
- RUDNICK, MARK; STERRITT, GRAHAM M.; and FLAX, MORTON. "Auditory and Visual Rhythm Perception and Reading Ability." *Child Development* 38: 581-87; June 1967.
- SAPIR, SELMA, and WILSON, BERNICE. "A Developmental Scale To Assist in the Prevention of Learning Disability." *Educational and Psychological Measurement* 27: 1061-68; Winter 1967.
- SILVER, ARCHIE A., and HAGIN, ROSA A. "Maturation of Perceptual Functions in Children with Specific Reading Disability." *Reading Teacher* 19: 253-59; January 1966.
- SILVER, ARCHIE A.; HAGIN, ROSA A.; and HERSH, MARILYN F. "Reading Disability: Teaching Through Stimulation of Deficit Perceptual Areas." *American Journal of Orthopsychiatry* 37: 744-52; July 1967.
- SMITH, DONALD E. P., and CARRIGAN, PATRICIA M. *The Nature of Reading Disability*. New York: Harcourt, Brace and Co., 1959. 149 pp.
- STERRITT, GRAHAM, and RUDNICK, MARK. "Auditory and Visual Rhythm Perception in Relation to Reading Ability in Fourth Grade Boys." *Perceptual and Motor Skills* 22: 859-64; June 1966.
- TAUBER, ROSALYN. "Identification of Potential Learning Disabilities." *Academic Therapy Quarterly* 2: 116-19, 123; Winter 1966-67.
- WILHELM, ROWENA. "Diagnostic Value of Test Score Differentials Found Between Measures of Visual and Auditory Memory in Severely Disabled Readers." *Academic Therapy Quarterly* 2: 42-44; Fall 1966.
- WILSON, MARGARET F. "Clinical Teaching and Dyslexia." *Reading Teacher* 21: 730-33; May 1968.

## CHAPTER VI

### Behavior Disorders

JOHN P. GLAVIN and HERBERT C. QUAY\*

Numerous publications have appeared recently on the subject of emotionally and socially maladjusted children. The majority of the articles, however, have been descriptions of projects, clinical case studies, or suggested methodologies without supporting data. For reasons of both space and scientific relevance, this review is limited mainly to research papers and articles from more readily available sources. As a result, theoretical approaches which lend themselves to programmatic research may be over-represented. In this report, learning theory, as noted by Balow (1966), continues to be a dominant force in the field. The few exceptions to the above limitations are primarily early reports of long-term studies which appeared to have major significance for the field.

Public school programs and teacher training facilities for emotionally and socially maladjusted children have flourished. Scheuer's (1966) survey of the United States and its affiliations showed that since 1962 there had been a three-fold increase in the number of states and U.S.-associated territories reporting special teacher training programs for disturbed children. A number of compilations of readings have appeared in recent years to accompany the growth of the field. Hellmuth's (1966) book consists of 19 papers covering a diverse range of topics, programs, and problems. Considerable emphasis is given to psychoeducational programs, perceptual development, and educational intervention technique. Long, Morse, and Newman's (1965) compilation contains a unique introductory section which deals with the experience of emotional disturbance as described by fiction writers. This is followed by an attempted consolidation of many classic articles of both a theoretical and a practical nature. Both books, particularly Hellmuth's, would have benefited by a careful analysis and discussion of the theoretical postures underlying the different articles.

Kessler (1966) united her own clinical experience with a broad review of the research and theoretical literature to produce a book which could serve as an excellent text or reference for those students in the field of childhood disturbances who prefer a psychoanalytic orientation. It is unfortunate that her last chapter which pertains to primary and secondary prevention did not lead to a reformulation of her earlier chapters. Her book would have had greater relevance to teachers of disturbed children if she

\*Acknowledgement is made of the helpful assistance of Phyllis Menter and Peter Witt, advanced graduate scholars at the University of Illinois. The preparation of this paper was supported in part by a grant (OEG-3-6-062063-1559) from the Office of Education, U.S. Department of Health, Education, and Welfare.



had dealt more fully with the recent trend toward educational intervention. On the whole, the above books have greatly enhanced the selection of textbooks in this area.

## Prevalence

### Screening and Predictive Measures

A major criterion for studies of maladjustment has been the use of teacher judgments. Using the basic screening techniques developed earlier by Bower, Maes (1966) did a multiple regression analysis on the significant variables in Bower's study. The variables, in decreasing order of prediction of clinical judgments, were (a) teacher rating (behavior), (b) arithmetic achievement, (c) intelligence, (d) a class play, (e) teacher rating (physical), and (f) reading achievement.

Westman, Rice, and Bermann (1967) collected nursery school records and elementary and high school cumulative records for the same children. An evaluation was made to see whether children who later show clinical signs of maladjustment could be identified during the nursery school years by making simple clinical judgments about their behavior. They concluded that observations by nursery school teachers were sufficiently accurate to place them in a key position to evaluate children's mental health. For predicting later adjustment, the most significant nursery school items observed by the teachers were relations with peers, eccentric behavior, deviant family structure, and pathological family reactions. However, the data from Rutter's (1967) study indicated that teacher evaluations alone cannot be regarded as valid indices of maladjustment.

Several studies attempted to develop predictive measures of later maladjustment. Pimm and McClure (1967) developed a checklist of specific, observable behaviors in the early grades which they believed to be predictive of later maladjustment. Their report revealed significant differences at the 0.001 level between high-risk and low-risk maladjustment groups on two of the criteria: referral to special services and current teacher rating of present classroom adjustment. Pate and Webb (1965) designed the Problem Screening Test as an untimed group test to select those children who would not make sufficient progress in the first grade to prepare them for a successful second year experience. The test sampled three areas of functioning: intellectual, emotional, and neurological. The total score identified 75 percent of the children who were not ready for second grade work, while one out of ten children was incorrectly identified.

Studies by Werry and Quay (1968) and Stevenson and his co-workers (1966) had important implications for the total screening process. Werry and Quay described a technique for measuring the classroom adjustment of elementary school children by means of direct frequency counts of certain deviant and work-related behaviors and teacher-pupil interactions.

The authors reported that the technique had value as a dependent variable measure of particular therapeutic interventions administered in the classroom setting and discriminated between normal and disturbed children. Stevenson and his colleagues studied whether adults would rate children of the opposite sex more positively than children of the same sex, a question which has important implications concerning the efficacy of teacher ratings in screening for emotional disturbance. The authors found no evidence to support the hypothesis of cross-sex effects, although the ratings appeared to indicate that adults rated girls in a more favorable manner than boys. The fact that the subjects came from a different academic background than most teachers and were untrained raters in observing children's behavior makes generalization of the findings tenuous.

The studies reviewed in this section suggested that the widespread use of teacher judgment in screening for disturbed children is generally justifiable; however, teacher estimations alone were not found to be a valid index of disturbance. Other inventories were reviewed which showed promise of being effective as screening devices in combination with teacher judgment. Since most global screenings of disturbed children are intended to institute some type of educational intervention program, it would seem appropriate that children initially screened as disturbed should be assessed in terms of classroom-relevant observable behavior before any educational intervention decisions are made. Furthermore, as Quay (1966) pointed out, the pattern of deviant behavior manifested by the child may have important implications for the particular educational strategy to be employed.

*Additional References:* Cary and Reveal (1967); Christiansen (1967); Feldhusen, Thurston, and Benning (1966).

### **Epidemiology in the Public Schools**

Since there was no widely accepted definition of maladjusted children, the figures from surveys on incidence of emotional disturbance in the public schools varied according to the definition and cutoff point employed by the researcher. The three following surveys were all made in rural, small-town areas of the country. A survey by Glavin (1968) in northeastern Tennessee indicated that 12.9 percent of the children could be classified as emotionally disturbed. Stennett (1966) reported that, of a northern Minnesota school district, 22 percent were screened out as either moderately or seriously emotionally handicapped. Both studies used a modified version of the teacher-peer-self identification process developed by Bower. McCaffrey and Cumming (1967) obtained their information from confidential interviews with teachers. Their first screening in New York found 4.2 percent of the pupils screened as emotionally disturbed.

Perhaps greater significance should be attached to some patterns of maladjustment which emerged from the above studies than to the findings concerning the incidence of disturbance. Stennett's finding of equal sex representation among the emotionally handicapped on his initial screening

was perhaps unique in the literature. McCaffrey and Cumming's results conformed to past studies, with approximately twice as many boys as girls being designated as emotionally disturbed. Concerning IQ, Stennett stated that the emotionally handicapped tended to make poorer scores than their classmates on group intelligence tests. Glavin reported that a group of children described as mentally retarded by their regular classroom teachers scored significantly lower in adjustment on the second screening and had a 36 percent incidence rate of disturbance. In addition, difficulty with academic learning appeared to be a major factor in emotional disturbance. Stennett mentioned that 78 percent of the boys and 66 percent of the girls had learning problems. McCaffrey and Cumming reported that nonpromotion and performance below class average were nearly twice as common among the disturbed children as among the total group and that approximately 66 percent of the boys and 50 percent of the girls classified as emotionally disturbed were reported as having learning problems.

### **Persistence of Emotional Disturbance**

Although nontreated control groups would improve experimental designs, their use in the evaluation of continuity and intervention studies has been limited. Aside from the questionable use of waiting-list or therapy-dropout control groups, nontreated controls have never been used in research on clinic populations. Most current continuity and intervention research has consisted of comparisons of the pre- and post-therapy status of a single treatment group or comparison of a treatment group with a waiting-list or therapy-dropout group. The failure to include nontreated controls in these studies has left crucial questions unanswered, such as the uncertain influence of the promise of future treatment on the waiting-list child's status or the equivalence of the dropout controls to the treatment group.

There have been studies recently reported which have circumvented many of the above-mentioned drawbacks. These results suggested that the majority of disturbed children improved without intervention. The design used by McCaffrey and Cumming (1967) involved three teacher identification surveys at two-year intervals. Children labeled as disturbed by only one rater comprised 70 percent of the boys and 73 percent of the girls. While their study may be criticized for relying solely upon teacher identification, McCaffrey and Cumming made a unique contribution to the field with their investigation of the behavior patterns associated with persistent disturbances. A major finding was the highly positive relationship between increased academic performance and improved emotional health. Shephard, Oppenheim, and Mitchell (1966) reported a comparison between a group of 50 children attending child guidance clinics and a group of supposedly healthy, nontreated children matched by age, sex, and degree of deviant behavior as determined by a clinical team. The results indicated that referral to a child guidance clinic was related as much to parental reactions (anxious, easily upset by stress, and lacking ability to cope with their



children) as to morbidity in the children. Approximately two-thirds of both groups improved over a two-year period. Unfortunately, the researchers relied solely upon a parents questionnaire and interview for their evaluations.

Glavin (1968) revealed improvement after a four-year interval in 70 percent of the children initially classified as disturbed who had not received any type of intervention. The "spontaneous" improvements were partially attributed to a statistical regression effect.

In contrast to the above studies, several reports concluded that a significant number of children identified as emotionally handicapped were not likely to resolve their adjustment problem without help. Stennett (1966) found an 85 percent persistence rate of emotional handicaps after one year and a 40 percent persistence rate after three years. He minimized the latter finding as due to an abridged technique used for his last screening.

In comparing gross nursery school adjustment ratings with later use of mental health services, Westman, Rice, and Berman (1967) found that 78 percent of the 22 children who showed a low early adjustment rating later sought the use of mental health services. In order to determine whether later problems did, in fact, reflect residual adjustment problems, it would have been helpful if the authors had stated the type of childhood disturbance and the purpose for which mental health services were sought in later life. In addition, a sample of 22 children was very small for use in this type of study. The subjects, predominantly of middle class socioeconomic backgrounds with many professional and university families, were a biased sample, more likely to seek therapy. Nevertheless, their correlation of 0.88 between low high school adjustment and the use of mental health services again directed attention to the strong relationship between academic and emotional adjustment.

Robins (1966) used the global responses of two psychiatrists to compare the adult social and psychiatric outcomes of 524 child guidance clinic patients with those of 100 normal schoolchildren. A 30-year follow-up after initial referral showed the antisocial children to be more often psychiatrically ill, with sociopathic personality as the most common diagnostic syndrome. In contrast, the neurotic children as adults resembled the control subjects. A major criticism in evaluating this study is that the normal control group might be considered supernormals. For example, 31 percent of the control group had IQ scores of 110 or better, in contrast to 16 percent of clinic patients. Secondly, a stringent educational criterion imposed for inclusion in the control group was perhaps a biasing factor. The control group contained no children who had repeated a grade, had been expelled, or had been transferred to a correctional institution in elementary school. Thirty-six percent of the control subjects had completed high school or had gone beyond a high school education while only 14 percent of the clinic group had done so. Despite these shortcomings, Robins' study was noteworthy for questioning many commonly held beliefs and for examin-

ing in detail the childhood factors that appeared to predict which disturbed children would grow into sociopathic adults.

The results of these persistence studies should not be confused with persistence rates for profoundly disturbed children. The reviewed studies referred only to children who remained in regular classes of the public schools. While the studies did apply to the vast majority of emotionally disturbed children who usually remained in the regular classes of the public schools, the work of Menolascino and Eaton (1967) with mentally retarded children and Rutter (1965) with primarily autistic children has suggested a much poorer prognosis for the relatively rare psychotic child.

Clarizio (1968) described many of the problems encountered in studying the persistence of emotional disturbance. He reviewed retrospective and follow-up studies and concluded that improvement seemed to characterize the course of disturbances in children as much as, or more than, chronicity or persistence. The conflicting studies on persistence generally support Clarizio's conclusion if one accepts Stennett's last screening at face value. Yet, despite its shortcomings, Robin's study appears to be the most rigorous methodologically.

## Characteristics

### School Achievement

Wolf (1965) compared children, diagnosed as functionally emotionally disturbed, with their siblings. No significant differences in scholastic grades were observed although the disturbed group received significantly lower effort grades. However, Wolf included only neurotic and personality-disordered children and excluded all children who were judged either as mentally retarded, brain damaged, psychotic, or as slow learners in special classes. In addition, the children were matched on a group IQ test. Since previous research has indicated that emotionally disturbed children tend to receive depressed scores with group IQ tests, Wolf's disturbed group actually had a higher mean intelligence than their sibling controls.

Schroeder (1965) revealed another possible basis for ambiguity on disturbed children's school achievement. She classified children into five categories: (a) those with psychosomatic problems, (b) those exhibiting aggressive behavior, (c) those having school difficulties, (d) those having school phobia, and (e) those having neurotic-psychotic personalities. Her data showed that educational disabilities were frequently concomitant with emotional disturbance. However, she concluded that disturbed children were not a homogeneous group with respect to academic performance. The school-difficulties children had the lowest mean achievement in arithmetic and reading, while the highest achievement was found among neurotic-psychotic children. All categories scored consistently lower in arithmetic than in reading. Motto and Lathan (1966) clarified some of the discrepancies in past research on academic achievement which used

varying periods of residential treatment. Children who remained in a hospital school in excess of 10 months benefited from their educational experience at the level expected for their mental ages, while those who attended for shorter periods were below the expected level.

Three studies were reported from which inferences may be drawn about the relationship between emotional disturbance and school achievement. Luszki and Schmuck (1965) reported improved functioning in school and greater self-esteem for elementary school students who perceived their parents' attitude toward school as supportive rather than nonsupportive. Brodie and Winterbottom (1967) compared 11 elementary school children, diagnosed as having psychogenic learning disturbances, with a control group of normal children. Based on blind clinical interviews with parents, the authors evaluated psychodynamic hypotheses and found that the learning-problem group was subject to more traumatic events and a more traumatic milieu. This group experienced more derogatory treatment and secrecy from their mothers than did the control group. Unfortunately, four of the mothers of learning-problem children had some casework treatment prior to the interview. Since a small sample was used for each group ( $N = 11$ ), one may question the hypothesis regarding secrecy and derogation. Graubard (1967) reported that the psycholinguistic correlates of reading disability in disturbed, delinquent children were poorer than those of normal children in the discrimination areas of visual motor association, auditory vocal automatic, visual motor sequential, mazes, and right-left discrimination. He proposed that there were deficits of the visual motor channel at the integrational level and presented suggestions concerning special remedial procedures.

It was satisfying to note that several of the above investigators had controlled for a number of relevant variables such as IQ, social class, and milieu in studying academic achievement. Length of stay in an institution and the type of disturbance have been shown to be additional pertinent variables. Nevertheless, as Quay, Morse, and Cutler (1966) have stated, the basic problem remains that children with different patterns of maladaptive behavior may well show different patterns of academic correlates. The tendency of most investigators to use either a global definition of emotional disturbance or poorly defined clinical categories precludes the systematic study of the relationships of academic achievement to the statistically derived dimensions of behavior disorders.

*Additional Reference:* Rautio (1966)

### **Social and Psychological Correlates**

Recent studies on the correlates of emotional and social maladjustment have focused on the relationship of the child to his parents and especially to his mother. Weinstein (1965) and Fisher (1967) compared the social schema of normal children with the schema of children with behavior problems by adapting Kuethe's technique for measuring social schemata.



The children were asked to place adult and childlike figures on a piece of paper. The physical distance between the figures was believed to be suggestive of the emotional distance that the child perceived in his relationship to his parents. Both Weinstein and Fisher found a tendency among normal elementary school boys to place child figures closer to adult female figures than to adult male figures. In contrast, emotionally disturbed boys placed child figures farther from adult female figures than from either adult male or peer figures. Fisher found that children who placed the child and adult female figures relatively far apart tended to have angry, hostile mothers. A third study by Gerber (1968) corroborated the finding that disturbed boys, as well as boys having learning disabilities, placed the dolls representing themselves and their mothers at greater distances than did normal male children. Parents of disturbed children made frequent use of a family grouping schema which isolated one doll from the other three. Additional studies are needed to determine whether other types of disturbed children and their parents have similar perceptions of their family relationships.

Other studies using objective tests offered support for these conclusions. Todd (1966) suggested that there was a measurable incompatibility of psychological needs between a mother and the child she considered emotionally disturbed. This incompatibility was manifested primarily in aggressive attitudes. Krug (1965) reported that clinic parents were found to be hostile, to show negative attitudes toward each other, and to be strict in their disciplinary methods as compared to parents of normal children. Differences between parents of conduct problem (acting-out) children and personality problem (anxious-withdrawn) children showed that the parents of conduct problem children were more hostile.

Of the three major indices of psychiatric illness, i.e., symptoms, impairment of functioning, and psychiatric referral, the latter has been studied the least thoroughly. Psychiatric referral was studied by Shephard, Oppenheim, and Mitchell (1966), but the focus was on background factors which distinguished the treated and nontreated groups matched for age, sex, and symptoms. Wolff (1967) studied the behavioral difficulties which occurred in children who alarmed or irritated the parents to the point of seeking psychiatric help. Using an interview with the mother and a behavior inventory completed by the teacher, Wolff found that clinic attenders were more discontented, had more soiling and enuresis, and displayed more antisocial behavior than nonreferred children. Psychiatric referral, in contrast to the results of Shephard, Oppenheim, and Mitchell, did not depend solely upon parental anxiety and intolerance of the child's symptoms.

Two final studies should be noted here since they focused on the significance of the adult-child relationship with respect to therapeutic intervention. D'Angelo and Walsh (1967) explored the efficacy of various approaches to short-term treatment. Improvement of the child was measured by psychodiagnostic evaluation and a questionnaire completed by the

teachers. Group I parents and children received no treatment, yet the child showed some improvement. This improvement, however, was not statistically significant. Group II, composed of children who were seen in individual therapy, demonstrated a significant decline in adjustment. Children whose parents received group therapy constituted Group III, and the children exhibited a significant positive trend in improvement. Children in Group IV, using the traditional approach with the child in individual therapy and the parent in group therapy, showed no change. Lisle (1968) randomly assigned pupils, judged by their teachers to be among the most poorly adjusted, to seven separate counseling groups. The counseling groups consisted of either the children, parents, and teachers, or some combination of the above. Lisle found no specific approach or combination of approaches which were superior to all the others for improving pupil adjustment as perceived by the pupil himself, his teacher, and his peers. However, those treatment approaches without pupil involvement were more effective for improving teacher-perceived pupil adjustment.

Replication of these studies with a larger sample size, better criteria measures, and more explicit description of procedures and diagnostic categories would be most welcome. In spite of these shortcomings, the studies appeared to support the proposition that improvement was more likely if therapy involved the parent rather than the disturbed child. This finding was not unexpected when the other studies reported in this section were considered. Whether by projective or objective testing, the results emphasized the mutual interrelationship between the child and his parents.

*Additional References:* Alexander and Leaverton (1967); Cowen and others (1965); Graubard (1967); Jenkins, NurEddin, and Shapiro (1966); Oakes (1966); Scarpitti (1965); Schrager and others (1966).

## Therapeutic Educational Provisions

Educational approaches such as teacher assistants, crisis or helping teachers, and the recently popular concept of the resource room have all been described at length in the literature, but no substantive research has yet been reported. Since these approaches have significant implications in terms of alleviating the manpower shortage, it is hoped that researchers in the future will focus on their efficacy. The meager research published on therapeutic camping continued to suffer from very small samples, questionable evaluations of emotional disturbance, and poor methods for assessing results.

*Additional References:* Connor and Muldoon (1967); McCreary-Juhasz and Jensen (1968); Morse (1966); Rickard and Dinoff (1967); Zax and others (1966).

## Models for Learning and Behavior Changes

A number of ongoing projects evaluating structured or reinforcement classroom paradigms were described in the literature, but only preliminary results were, as yet, available. Hewett (1967) developed and used an engineered classroom design based on the behavior modification model in both institutional and public schools. The engineered classroom attempted to provide a setting for implementation of a hierarchy of educational tasks, meaningful rewards for learning, and an appropriate degree of teacher structure. Hewett, Taylor, and Artuso (1968) reported that children in the engineered classrooms maintained a 5-20 percent task-attention advantage over children in the control classrooms. Gains in arithmetic fundamentals were significantly correlated with the use of the engineered design, but reading and spelling gains were not significantly different between the experimental and control conditions.

An innovative pilot study by Minuchin, Chamberlain, and Graubard (1967) presented an overview of socialization processes in disorganized, low socioeconomic families followed by some assumptions about the influence of these processes on the learning style of the child. They devised an intervention curriculum designed to explore ways of enabling six delinquent children to overcome their deficits. The pupils alternately assumed a role as participant or observer and were coached to rate the other children's ability to respond to the teacher and to enumerate behaviors which enhanced or interfered with learning. Small monetary rewards were given for points received in the classroom with the judges also being rated and given points. Based on the authors' clinical observations of the children, marked increases were noted in ability to maintain attention and in mechanics of dialogue and language.

Lewis (1967) reported on the early follow-up results of Project Re-Ed. The experimental pattern required two specially trained teachers who lived and worked with eight children in a round-the-clock residential program which averaged six months per child. The emphasis was on education and the ecological modification of the child's natural environment. The designation most frequently given the children was "adjustment reaction of childhood." Evaluations of the children before and at the end of treatment as well as follow-up evaluations by the child's parents, teachers, and peers agreed, within fairly narrow limits, that between 75 percent and 87 percent of the children were functioning much better upon graduation from Re-Ed than when they were enrolled. In addition to the excellent improvement rate, the Re-Ed results deserved careful study since less time and expense were spent in treatment than is customary. It is hoped that future reports will include a breakdown of results by the type of disorder as well as by the relationship between the type of problem and the ecological modifications most successful in dealing with it.

*Additional References:* Phillips, El-Batrawi, and Tanck (1966); Rhodes (1967).



## The Effects of Special Class Placement

Scheuer (1966) noted that the increasing demand of educational provisions for the emotionally disturbed has placed a greater reliance upon the teacher as the primary intervention agent. In many cases the teacher has been ill prepared to assume this role. The usual pattern has been to place a teacher in charge of a self-contained special class similar to that for other types of exceptional children. Recently, one study has reported on how disturbed children viewed being assigned to a special class, and several studies have attempted to evaluate the academic progress of disturbed pupils in special class programs. Whelan (1966b) found that the semantic differential could discriminate between groups of normal and disturbed school-age males. The emotionally disturbed children in special classes and the normal class children attached more positive meanings to school, self, and family-related concepts than did the emotionally disturbed regular class children who were selected from a waiting list of a community child guidance clinic.

Safford and Watts (1967) evaluated the academic progress of 27 disturbed pupils in special classes, diagnosed as having minimal central nervous system (CNS) dysfunction. They found that these pupils were significantly below grade level when they entered the program and that the children tended to grow at one-third the expected rate in academic areas while in the special class. No descriptive data were given nor was there any information concerning the teaching methods, classroom management techniques, specific learning disabilities, or types of disturbance. Radin and others (1966) expanded on the traditional orthopsychiatric team to include special educators who taught one primary and one intermediate special class. Their study demonstrated an academic gain of 2.4 years on the *California Achievement Test* as compared with a gain of 0.9 years for a control sample of disturbed children in a normal public school program. Because of the selective criteria used for admission to the program, it was doubtful that this group of children was representative of disturbed children in general. The child had to have at least a normal IQ, no CNS dysfunction, and be able to establish and maintain group relations. The parents had to be willing to participate in counseling. Camp and Lathen (1967) utilized a trained teacher in a special class with a maximum of eight children. Although the authors commented that the teacher was permissive or semistructured, the description of the program seemed quite structured. An average academic gain of nine months was reported. The children's previous performance in regular classes had been six months per year. While the authors presented a good description of their program rationale, their article was more of a clinical than a scientific study, with no evidenced control of variables and little data presented.

Rubin, Simson, and Betwee (1966) employed an excellent research design to test the efficacy of special class placement. While the results showed a few propitious indications for special classes, the major con-

clusion was that special classes did not produce results significantly different from regular classes. Mean academic gains for both groups showed less than the expected one-year increment per year. A description of the type of classroom philosophy, teaching style, and educational methods would have been most helpful. Furthermore, although the children were classified into four categories of disturbance, it seemed that they were randomly grouped together for comparable intervention experiences.

While support was given in the research literature to the special class approach for promoting the child's improved feelings, the evidence on academic achievement was conflicting. Moreover, the studies on achievement suffered from biased samples, poor methodology, and very poor design, aside from the sound research methodology employed by Rubin, Simson, and Betwee.

### Prevention and School Consultation

Both educators and clinicians have long complained of the difficulties in communicating with each other. Recently there have appeared several descriptive articles on consultation services for normal teachers of disturbed children but only one study even remotely comparable to the consultation approach presented supporting data.

Cowen and his colleagues (1966) reported on a cross validation of their earlier study on the effectiveness of a preventive program and the sequelae of early-detected pathology. The prevention program was applied in the experimental school and consisted of early diagnostic evaluation, social work interviews with the mothers, consultative services, after-school activity programs, and parent-teacher discussion groups. They reported that 7 of the 20 school and adjustment criteria differentiated between the experimental and the control groups, all in favor of the experimental group. After three years, those children identified as disturbed were less favorable on 19 of the 20 criteria, 14 of them significant at the 0.05 level. The authors concluded that disturbed children seemed to be on a global downhill course by third grade.

In spite of methodological shortcomings such as subjective criteria for evaluation and a poor description of the control group, the findings of Cowen and his associates point to the valuable potential of a community-based preventive approach, not necessarily restricted to the school setting. Future research might wish to compare disturbed children with and without a preventive program.

*Additional References:* Brown (1967); Rosenblum and Ottenstein (1965); Seagull and Johnson (1966).

### Approaches to Behavioral Change and Management

In addition to excellent theoretical presentations of the relevance of learning theory procedures for teachers of disturbed children, the literature

suggested a trend away from the traditional procedure of treating the individual child to the more comprehensive application of behavior modification principles to the entire classroom. Whelan (1966a) reviewed behavior modification procedures in terms of their relevance for classroom teachers. Of particular note was the author's clarification of the relationship of behavior modification procedures to other types of intervention approaches and his discrimination between punishment and deceleration procedures. Clarizio and Yelon (1967) discussed the potential advantages for teachers of utilizing the techniques of extinction, positive reinforcement, modeling, punishment, discrimination learning, and desensitization. They emphasized that these techniques could be taught to teachers through laboratory courses taken in conjunction with their formal course work or through in-service meetings and workshops. Quay and his associates (1966) offered a conceptualization of the nature of children's behavior disorders in relationship to principles of behavior modification. They later reported (Quay and others, 1967) an experiment in which the attending behavior of conduct problem pupils rose from 41 percent during baseline to 61 percent for the entire reinforcement period. In the last third of the reinforcement period, attending behavior was 76 percent.

Several projects described token or social reinforcement systems which could be used by one teacher in an average classroom. O'Leary and Becker (1967), using a token reinforcement program, obtained an abrupt reduction in deviant behavior for the eight most disruptive children in an adjustment class. Delay of reinforcement was gradually increased to four days without a concomitant increase in deviant behavior. Of particular significance was the authors' report that the children's appropriate behavior generalized to other school situations. McKenzie and his associates (1968) used naturalistic classroom events capable of providing incentives as consequences for academic progress. However, when the children's academic behaviors failed to reach an optimal level, their parents agreed to supply the children's allowances contingent upon their weekly grades. Significant increases in the children's academic behavior were reported.

These articles suggested an increased likelihood of adoption by the classroom teacher of behavior modification techniques. The development of procedures readily teachable and applicable to group settings has enhanced the utility of behavior modification principles for educators.

*Additional References:* Allen and others (1967); Becker and others (1967); Blom (1966); Brown and Shields (1967); Burchard and Tyler (1965).

## Teacher Behavior, Training, and Role

Several studies and theoretical articles have been reported on the current status of teacher certification, training, and the classroom behavior styles of teachers of the emotionally disturbed. A number of reports dealt with the behavior styles of teachers of the disturbed both in regular and in



special classes. Johnson (1965) collected data with the Teacher Preference Schedule, a test presumed to measure unconscious motives, and the Theoretical Orientation to Teaching the Emotionally Disturbed Child Questionnaire, also thought to measure unconscious motives, educational framework, and classroom approach. Discernible differences were found to exist between teacher behavior styles within and between groups drawn from five types of treatment centers. The child-permissive pattern was found to be the most dominant while the teacher-centered, structured pattern was second. The permissive style focused on high nondirectiveness and nurturance combined with low dominance and orderliness. Both male and female teachers scored high on nurturance and nondirectiveness, but males tended to be more organized and self-assured.

Kounin, Friesen, and Norton (1966) found that, although disturbed children manifested less school-appropriate behavior than the nondisturbed children, both types of children showed similar kinds of behavior changes in varying regular classroom settings. Teachers who were successful in managing the behavior of nondisturbed children were also relatively successful with disturbed children. The degree to which the teacher was able to communicate to the class the impression that she knew what was occurring, her ability to program for variety in learning activities, and her techniques of handling group movement and transitions were all positively correlated with improved children's behavior. One of the most important implications of this study was that concrete, readily teachable techniques could be delineated for facilitating improved classroom behavior. The future training of education students in the techniques of group management was strongly suggested. A later study by Kounin and Obradovic (1968) replicated and extended the earlier study using better work samples. Additional follow-ups and similar research were still needed, with greater delineation of the types of disturbed children studied, teacher personality factors, and classroom philosophy and procedures.

Two studies attempted to determine factors associated with teachers' perception of emotional disturbance and the reasons given for preferring to teach the disturbed child. Dobson (1966) found that teachers of middle class children differed significantly from those of culturally deprived children in their increased perception of the seriousness of behavioral problems. Teachers with 10 years experience from both groups viewed undesirable acts as being less serious than those teachers with 3-10 years experience. No significant differences in attitudes toward treatment of behavioral problems existed between teachers of middle class and culturally deprived children. Jones and Gottfried (1966), administering the *Edwards Personal Preference Schedule* and the *Teacher Preference Schedule*, found a large number of persons not currently teaching or preparing to teach exceptional children who expressed interest in teaching gifted, emotionally disturbed, or delinquent children. Another interesting finding, in light of Johnson's (1965) results, was Jones and Gottfried's report of

high scores on the nurturance and nondirective subscales which differentiated high from low degree of preference for teaching delinquents and emotionally disturbed children. The authors cautioned, however, that it would be premature to make decisive statements about specific patterns of psychological needs allied to preferences for teaching these children.

Some descriptive and theoretical articles pertaining to the training and role of teachers for disturbed children are cited below. It is a pleasure to report that these impressions and viewpoints on "what ought to be" are now being supplemented by long-needed research in an area grossly overlooked in the past. The above articles have described the rapid increase in demand for teachers of disturbed children and have cited an apparent pool of untapped educators and educational students who are potentially interested in teaching in this field. The training of these future teachers should certainly be affected by the excellent studies of Kounin and his associates (1966). They have indicated that a teacher skilled in group management techniques may often be able to contain the disturbed child either in the regular or special classroom and that these management techniques are readily teachable.

*Additional References:* Amidon (1967); Balow (1966); Beck (1966); Cotter (1966); Hewett (1966); Johnson (1968)

### Prospective and Suggested Research

The increase in special classes for the emotionally or socially maladjusted has led to an increased emphasis upon use of educationally based interventions. While research on classroom behavior styles of teachers suggested that teachable techniques can be delineated for facilitating improved classroom behavior, the research also suggested that with these techniques the disturbed child could often be managed in the regular, as well as the special, class. Furthermore, research on the effects of special class placement has been conflicting.

There simply is not a sufficient number of trained teachers for the special classes now operating, much less for those anticipated in the future. This manpower shortage is potentially less acute in the field of educating the emotionally disturbed than in other educational fields, since there is apparently a large number of persons not currently teaching or preparing to teach emotionally or socially maladjusted children who have expressed an interest in this area. The problem is to tap this potential resource by some innovation such as beginning training for special class teachers early in their undergraduate program and having a number of increasingly difficult practicum experiences which perhaps could include salaried positions.

Since we have not yet reached the point where we can confidently distinguish between a child whose disturbance will persist and one who

will "spontaneously" improve, the emphasis for educational intervention perhaps should be placed on eliminating concrete observable behavior of immediate concern and remediating the frequently concomitant academic learning problems rather than on the more ambitious restructuring of personality. Of particular importance for educators, recent learning theory research has shifted from the use of conditioning to influence an individual child's behavior to the application of operant principles to the entire classroom. Whatever the emphasis in future placement policies, it would seem that if an educational intervention is to be most efficiently utilized, the referral should be based on the child's specified behavioral difficulties, which will often be accompanied by learning deficits, rather than on the vague label that he is emotionally disturbed.

The research showing the importance of parental influences upon the child suggested the need for a more encompassing intervention program rather than reliance upon a single point of attack. Application of a modified version of the ecological model to the public schools and perhaps group therapy or training in behavioral management for the parents may be necessary. These approaches would require that the public schools view the disturbed child as being more than their customary nine-to-three o'clock responsibility. This broader concept of the accountability of special education in the socialization of disturbed children may well be necessary for many of the children now in our charge.

Many studies in this review appeared to offer different paths to successful intervention. Most experts would be likely to advocate some combination of models rather than a single approach. Perhaps it is pertinent to point out that the word *model* is used when another science provides the explanatory system—learning theory, psychoanalysis, ecology, and so forth—and to wonder at the relative lack of educational influence in the field. What is clearly needed is a comprehensive conceptual framework based on the principles of teaching and learning from which effective and efficient strategies for classroom management and the acquisition of academic skills logically follow.

## Bibliography

- BALOW, BRUCE. "The Emotionally and Socially Handicapped." *Review of Educational Research* 36: 120-33; February 1966.
- BRODIE, RICHARD D., and WINTERBOTTOM, MARIAN R. "Failure in Elementary School Boys as a Function of Traumata, Secrecy, and Derogation." *Child Development* 38: 701-11; September 1967.
- CAMP, WILLIAM L., and LATHAN, LEVI. "A Successful Classroom Program for Emotionally Disturbed Children." *Training School Bulletin* 64: 31-38; May 1967.
- CLARIZIO, HARVEY F., and YELON, STEVEN L. "Learning Theory Approaches to Classroom Management: Rationale and Intervention Techniques." *Journal of Special Education* 1: 267-74; Spring 1967.
- CLARIZIO, HARVEY F. "Stability of Deviant Behavior Through Time." *Mental Hygiene* 52: 288-93; April 1968.



- COWEN, EMORY L., and OTHERS. "Prevention of Emotional Disorders in the School Setting: A Further Investigation." *Journal of Consulting Psychology* 30: 381-87; October 1966.
- D'ANGELO, RITA Y., and WALSH, JOHN F. "An Evaluation of Various Therapy Approaches with Lower Socioeconomic Group Children." *Journal of Psychology* 67: 59-64; September 1967.
- DOBSON, RUSSELL LEE. *The Perception and Treatment by Teachers of the Behavioral Problems of Elementary School Children in Culturally Deprived and Middle-Class Neighborhoods*. Doctor's thesis. Norman: University of Oklahoma, 1966. 125 pp. Abstract: *Dissertation Abstracts* 27: 1702A-03A; No. 6, 1967.
- FISHER, RHODA LEE. "Social Schema of Normal and Disturbed School Children." *Journal of Educational Psychology* 58: 88-92; April 1967.
- GERBER, GWEN L. *Psychological Closeness Expressed Through Spatial Relationships by Families of Normal and Disturbed Children*. Doctor's thesis. Los Angeles: University of California, 1967. 146 pp. (Abstract: *Dissertation Abstracts* 28: 4756B; No. 11, 1968.
- GLAVIN, JOHN P. "Spontaneous" Improvement in Emotionally Disturbed Children. Doctor's thesis. Nashville, Tenn.: George Peabody College for Teachers, 1967. 133 pp. Abstract: *Dissertation Abstracts* 28: 3503A; No. 9, 1968.
- GRAUBARD, PAUL S. "Psycholinguistic Correlates of Reading Disability in Disturbed Delinquent Children." *Journal of Special Education* 1: 363-68; Summer 1967.
- HELLMUTH, JEROME, editor. *Educational Therapy*. Vol. 1. Seattle, Wash.: Special Child Publications, 1966. 468 pp.
- HEWETT, FRANK M. "Educational Engineering with Emotionally Disturbed Children." *Exceptional Children* 33: 459-67; March 1967.
- HEWETT, FRANK M.; TAYLOR, FRANK D.; and ARTUSO, ALFRED A. "The Santa Monica Project." *Exceptional Children* 34: 387; February 1968.
- JOHNSON, JOHN L. *Classroom Behavior Styles Reported by Teachers of Emotionally Disturbed Children*. Doctor's thesis. East Lansing: Michigan State University, 1965. 200 pp. Abstract: *Dissertation Abstracts* 26: 4450; No. 8, 1966.
- JONES, REGINALD L., and GOTTFRIED, NATHAN W. "Psychological Needs and Preferences for Teaching Exceptional Children." *Exceptional Children* 32: 313-21; January 1966.
- KESSLER, JANE W. *Psychopathology of Childhood*. Englewood Cliffs, N.J.: Prentice-Hall, 1966. 533 pp.
- KOUNIN, JACOB S., and OBRADOVIC, SYLVIA. "Managing Emotionally Disturbed Children in Regular Classrooms: A Replication and Extension." *Journal of Special Education* 2: 129-35; Winter 1968.
- KOUNIN, JACOB S.; FRIESEN, WALLACE V.; and NORTON, A. EVANGELINE. "Managing Emotionally Disturbed Children in Regular Classrooms." *Journal of Educational Psychology* 57: 1-13; February 1966.
- KRUG, RONALD S. *Attitudinal and Behavioral Differences in Childrearing Practices Between Parents of Clinic and Nonclinic Children: A Questionnaire Assessment*. Doctor's thesis. Urbana: University of Illinois, 1965. 181 pp. Abstract: *Dissertation Abstracts* 26: 2869-70; No. 5, 1965.
- LEWIS, WILBERT W. "Project Re-Ed: Educational Intervention in Discordant Child Rearing Systems." *Emergent Approaches to Mental Health Problems*. (Edited by Emory L. Cowen, Elmer A. Gardner, and Melvin Zax.) New York: Appleton-Century-Crofts, 1967. Chapter 19, pp. 352-68.
- LISLE, JAMES D. *The Comparative Effectiveness of Various Group Procedures Used with Elementary Pupils with Personal-Social Adjustment Problems*. Doctor's thesis. Kent, Ohio: Kent State University, 1967. 198 pp. Abstract: *Dissertation Abstracts* 28: 4485A; No. 11, 1968.
- LONG, NICHOLAS J.; MORSE, WILLIAM C.; and NEWMAN, RUTH G., editors. *Conflict in the Classroom: The Education of Emotionally Disturbed Children*. Belmont, Calif.: Wadsworth Publishing Co., 1965. 515 pp.
- LUSZKI, MARGARET B., and SCHMUCH, RICHARD. "Pupil Perceptions of Parental Attitude Toward School." *Mental Hygiene* 49: 296-307; April 1965.
- MCCAFFREY, ISABEL, and CUMMING, JOHN. *Behavior Patterns Associated with Persistent Emotional Disturbances of School Children in Regular Classes of Elementary Grades*. Onondaga County: Mental Health Research Unit, New York State Department of Mental Hygiene; December 1967. 23 pp.

- MCKENZIE, HUGH S., and OTHERS. "Behavior Modification of Children with Learning Disabilities Using Grades as Tokens and Allowances as Back-Up Reinforcers." *Exceptional Children* 34: 745-52; Summer 1968.
- MAES, WAYNE R. "The Identification of Emotionally Disturbed Elementary School Children." *Exceptional Children* 32: 607-609; May 1966.
- MENOLASCINO, FRANK J., and EATON, LOUISE. "Psychoses of Childhood: A Five Year Follow-Up Study of Experiences in a Mental Retardation Clinic." *American Journal of Mental Deficiency* 72: 370-80; November 1967.
- MINUCHIN, SALVADOR; CHAMBERLAIN, PAMELA; and GRAUBARD, PAUL. "A Project To Teach Learning Skills To Disturbed, Delinquent Children." *American Journal of Orthopsychiatry* 37: 558-67; April 1967.
- MOTTO, JOSEPH J., and LATHAN, LEVI. "An Analysis of Children's Educational Achievement and Related Variables in a State Psychiatric Hospital." *Exceptional Children* 32: 619-23; May 1966.
- O'LEARY, K. DANIEL, and BECKER, WESLEY C. "Behavior Modification of an Adjustment Class: A Token Reinforcement Program." *Exceptional Children* 33: 637-42; May 1967.
- PATE, JOHN E., and WEBB, WARREN W. "Screening Beginning First Graders for Potential Problems." *Exceptional Children* 32: 111; October 1965.
- PIMM, JUNE B., and McCLURE, GORDON. "A Screening Device for Early Detection of Emotional Disturbance in a Public School Setting." *Exceptional Children* 33: 647-48; May 1967.
- QUAY, HERBERT C. "Dimensions of Problem Behavior in Children and Their Interaction in the Approaches to Behavior Modifications." *Kansas Studies in Education* 16: 6-13; June 1966.
- QUAY, HERBERT C.; MORSE, WILLIAM C.; and CUTLER, RICHARD L. "Personality Patterns of Pupils in Special Classes for the Emotionally Disturbed." *Exceptional Children* 32: 297-301; January 1966.
- QUAY, HERBERT C., and OTHERS. "Remediation of the Conduct Problem Child in the Special Class Setting." *Exceptional Children* 32: 509-15; April 1966.
- QUAY, HERBERT C., and OTHERS. "Conditioning Visual Orientation of Conduct Problem Children in the Classroom." *Journal of Experimental Child Psychology* 5: 512-17; December 1967.
- RADIN, SHERWIN S., and OTHERS. "Orthopsychiatry and Special Services for Emotionally Disturbed Children in the Public School Setting." *Journal of School Health* 36: 245-48; June 1966.
- ROBINS, LEE N. *Deviant Children Grown Up*. Baltimore: Williams and Wilkins Co., 1966. 351 pp.
- RUBIN, ELI Z.; SIMSON, CLYDE B.; and BETWEE, MARCUS C. *Emotionally Handicapped Children and the Elementary School*. Detroit: Wayne State University Press, 1966. 286 pp.
- RUTTER, MICHAEL. "The Influence of Organic and Emotional Factors on the Origins, Nature, and Outcome of Childhood Psychosis." *Developmental Medicine and Child Neurology* 7: 518-28; October 1965.
- RUTTER, MICHAEL. "A Children's Behavior Questionnaire for Completion by Teachers: Preliminary Findings." *Journal of Child Psychology and Psychiatry and Allied Disciplines* 8: 1-11; No. 1, 1967.
- SAFFORD, ALTON L., and WATTS, CHARLES A. "An Evaluation of a Public School Program for Educationally Handicapped Children." *California Journal of Educational Research* 18: 125-32; May 1967.
- SCHUEER, ARNOLD L. "Certification, Teacher Preparation, and Special Classes for the Emotionally Disturbed and Socially Maladjusted: Report by States." *Exceptional Children* 33: 120-21; October 1966.
- SCHROEDER, LILY B. "A Study of the Relationships Between Five Descriptive Categories of Emotional Disturbance and Reading and Arithmetic Achievement." *Exceptional Children* 32: 111-12; October 1965.
- SHEPARD, MICHAEL; OPPENHEIM, A. N.; and MITCHELL, SHEILA. "Childhood Behavior Disorders and the Child-Guidance Clinic: An Epidemiological Study." *Journal of Child Psychology and Psychiatry* 7: 39-52, No. 1, 1966.
- STENNETT, R. G. "Emotional Handicap in the Elementary Years: Phase or Disease?" *American Journal of Orthopsychiatry* 36: 444-49; April 1966.

- STEVENSON, HAROLD W., and OTHERS. "Adult Ratings of Children's Behavior." *Child Development* 37: 929-41; December 1966.
- TODD, BEVERLEY C. *A Comparative Study of Mother-Child Relationships with Disturbed Children and Their Siblings*. Doctor's thesis. Philadelphia: Temple University, 1965. 146 pp. (Abstract: *Dissertation Abstracts* 26: 4819-20; No. 8, 1966.)
- WEINSTEIN, LAURA. "Social Schemata of Emotionally Disturbed Boys." *Journal of Abnormal Psychology* 70: 457-61; December 1965.
- WERRY, JOHN S., and QUAY, HERBERT C. "A Method of Observing Classroom Behavior of Emotionally Disturbed Children." *Exceptional Children* 34: 389; February 1968.
- WESTMAN, JACK C.; RICE, DALE L.; and BERMAN, ERIC. "Nursery School Behavior and Later School Adjustment." *American Journal of Orthopsychiatry* 37: 725-31; July 1967.
- WHELAN, RICHARD J. "The Relevance of Behavior Modification Procedures for Teachers of Emotionally Disturbed Children." *Intervention Approaches in Educating Emotionally Disturbed Children*. (Edited by Peter Knoblock.) Syracuse, N.Y.: Syracuse University Press, 1966. pp. 35-78. (a)
- WHELAN, RICHARD J. *Semantic Differential Behavior of Normal and Emotionally Disturbed School Age Males*. Doctor's thesis. Laurence: University of Kansas, 1966. 197 pp. (Abstract: *Dissertation Abstracts* 27: 1677A-78A; No. 6, 1966. (b))
- WOLF, MARTIN G. "Emotional Disturbance and School Achievement." *Journal of School Psychology* 4: 16-18; Autumn 1965.
- WOLFF, SULA. "Behavioral Characteristics of Primary School Children Referred to a Psychiatric Department." *British Journal of Psychiatry* 113: 885-93; August 1967.

### Additional References

- ALEXANDER, THIERON, and LEAVERTON, PAUL. "Differentiation Between Normal and Disordered Children by a Computer Analysis of Emotional and Verbal Behavior." *Journal of Psychology* 67: 141-46; September 1967.
- ALLEN, K. EILEEN, and OTHERS. "Control of Hyperactivity by Social Reinforcement of Attending Behavior." *Journal of Educational Psychology* 58: 231-37; August 1967.
- AMIDON, EDMUND. "An Approach to the Study of Teaching." *The Teaching-Learning Process in Educating Emotionally Disturbed Children*. (Edited by Peter Knoblock and John L. Johnson.) Syracuse, N.Y.: Syracuse University Press, 1967. pp. 7-26.
- BALOW, BRUCE. "A Program of Preparation for Teachers of Disturbed Children." *Exceptional Children* 32: 455-60; March 1966.
- BECK, GAYLE R. "Training of Teachers of the Emotionally Disturbed." *Journal of Teacher Education* 17: 137-41; Summer 1966.
- BECKER, WESLEY C., and OTHERS. "The Contingent Use of Teacher Attention and Praise in Reducing Classroom Behavior Problems." *Journal of Special Education* 1: 287-307; Spring 1967.
- BLOM, GASTON E. "Psychoeducational Aspects of Classroom Management." *Exceptional Children* 32: 377-83; February 1966.
- BROWN, EDWIN R., and SHIELDS, ELOISE. "Results with Systematic Suspension: A Guidance Technique To Help Children Develop Self-Control in Public School Classrooms." *Journal of Special Education* 1: 425-37; Summer 1967.
- BROWN, JONATHAN W. "Pragmatic Notes on Community Consultation with Agencies." *Community Mental Health Journal* 3: 399-405; Winter 1967.
- BURCHARD, JOHN, and TYLER, VERNON, JR. "The Modification of Delinquent Behavior Through Operant Conditioning." *Behavior Research and Therapy* 2: 245-50; April 1965.
- CARY, ARA C., and REVEAL, MARY T. "Prevention and Detection of Emotional Disturbance in Preschool Children." *American Journal of Orthopsychiatry* 37: 719-24; July 1967.
- CHRISTIENSEN, TED. "A Method of Identifying Maladjusted Children in the Classroom." *Mental Hygiene* 51: 574-75; October 1967.
- CONNOR, EILEEN M., and MULDOON, JOHN F. "Resource Programming for Emotionally Disturbed Teen-Agers." *Exceptional Children* 34: 261-65; December 1967.
- COTTER, KATHERINE C. "Emotionally Disturbed Children and Teacher Preparation." *Catholic Educational Review* 64: 457-69; October 1966.



- COWEN, EMORY L., and OTHERS. "The Relationship of Anxiety in School Children to School Record, Achievement, and Behavioral Measures." *Child Development* 36: 685-95; September 1965.
- FELDHUSEN, JOHN F.; THURSTON, JOHN R.; and BENNING, JAMES J. "Sentence Completion Responses and Classroom Social Behavior." *Personnel and Guidance Journal* 45: 165-70; October 1966.
- GRAUBARD, PAUL S. "The Use of Peabody Picture Vocabulary Test in the Prediction and Assessment of Reading Disability in Disturbed Children." *Journal of Educational Research* 61: 3-5; September 1967.
- HEWETT, FRANK M. "A Hierarchy of Competencies for Teachers of Emotionally Handicapped Children." *Exceptional Children* 33: 7-11; September 1966.
- JENKINS, RICHARD L.; NUREDDIN, ELHAM; and SHAPIRO, IRIS. "Children's Behavior Syndromes and Parental Responses." *Genetic Psychology Monographs* 74: 261-329; November 1966.
- JOHNSON, JOHN L. "Teacher Preparation for Educating the Disturbed: Graduate, Undergraduate, or Functional?" *Exceptional Children* 34: 345-51; January 1968.
- MCCREARY-JUHASZ, ANNE, and JENSEN, S. E. "Benefits of a School Camp Experience to Emotionally Disturbed Children in Regular Classrooms." *Exceptional Children* 34: 353-54; January 1968.
- MORSE, WILLIAM C. "Public Schools and the Disturbed Child." *Intervention Approaches in Educating Emotionally Disturbed Children*. (Edited by Peter Knoblock.) Syracuse, N.Y.: Syracuse University Press, 1966. pp. 113-28.
- OAKES, CHARLES G. *Parental Aspirations for Children That Are Associated with Emotional Disturbance in Young Sons*. Doctor's thesis. Atlanta: Emory University, 1966. 231 pp. Abstract: *Dissertation Abstracts* 27: 835A-36A; No. 3, 1966.
- PHILLIPS, E. LAKIN; EL-BATRAWI, SALAH; and TANCK, ROLAND H. "A School-Based Therapeutic Environment for Emotionally Disturbed Children." *High School Journal* 49: 254-58; March 1966.
- RAUTIO, EDWIN A. *A Comparison of Intelligence and Academic Achievement with Ratings on a Classroom Method for Screening Children for Emotional Handicaps*. Doctor's thesis. Lincoln: University of Nebraska, 1966. 101 pp. Abstract: *Dissertation Abstracts* 27: 602B; No. 2, 1966.
- RHODES, WILLIAM C. "The Disturbing Child: A Problem of Ecological Management." *Exceptional Children* 33: 449-55; March 1967.
- RICKARD, HENRY C., and DINOFF, MICHAEL. "Behavior Change in a Therapeutic Summer Camp: A Follow-Up Study." *Journal of Genetic Psychology* 110: 181-83; June 1967.
- ROSENBLUM, GERSHEN, and OTTENSTEIN, DONALD. "From Child Guidance to Community Mental Health: Problems in Transition." *Community Mental Health Journal* 1: 276-83; Fall 1965.
- SCARPITTI, FRANK R. "Delinquent and Nondelinquent Perceptions of Self, Values, and Opportunity." *Mental Hygiene* 49: 399-404; July 1965.
- SCHRAGER, JULES, and OTHERS. "The Hyperkinetic Child: Some Consensually Validated Behavioral Correlates." *Exceptional Children* 32: 635-37; May 1966.
- SEAGULL, ARTHUR A., and JOHNSON, JOHN L. "Second Stage Intervention: Reality Based Consultation for Teachers of the Emotionally Disturbed." *Intervention Approaches in Educating Emotionally Disturbed Children*. (Edited by Peter Knoblock.) Syracuse, N.Y.: Syracuse University Press, 1966. pp. 99-111.
- ZAX, MELVIN, and OTHERS. "A Teacher-Aide Program for Preventing Emotional Disturbance in Young Schoolchildren." *Mental Hygiene* 50: 406-15; July 1966.

## CHAPTER VII

### The Deaf and the Hard of Hearing

STEPHEN P. QUIGLEY

During the past several years there has been a considerable increase in the volume of research in the area of deafness. However, most of this research has been concerned with deaf children and very little with those who are hard of hearing. The hard of hearing group, which must represent a large population of handicapped children, appeared to be among the most neglected in terms of research and educational provisions. The increase in research relating to deaf children was largely confined to psychological studies and to technological efforts in the area of communication. With the possible exception of increased research on the use of manual communication, very little of the published work dealt directly with improving educational management of deaf children. With deaf adults the situation was somewhat different. There was an increased volume of studies relating directly to the provision of rehabilitation services, the major trend of which seemed to be the provision of community services for deaf persons on the same scale as presently exists for the general population.

One occurrence in recent years should be singled out for special note. This was the National Conference on the Education of the Deaf sponsored by the U.S. Department of Health, Education, and Welfare in 1967. The publication of the proceedings of the conference *Education of the Deaf: The Challenge and the Charge* (1968) could well provide a blueprint for much of the future research and educational planning for deaf children.

#### Incidence

##### General Population

Gentile, Schein, and Haase (1967) reported data on persons in the general population with binaural loss of hearing, classified in terms of ability to hear and understand speech. The data were based on information collected in household interviews of a representative sample of the U.S. population. Data were provided on the age at onset of hearing loss, utilization of hearing aids, special training and treatment received, and other pertinent factors. A little more than 4,000,000 persons were reported to have some loss of hearing in both ears. Among this group, 856,000 persons were classified as unable to hear and understand speech without the use of a hearing aid; 736,000 were classified as able to hear and understand

a few spoken words; and 2,439,000 were classified as able to hear and understand most spoken words. About 22 percent of the hearing-impaired population reported current use of hearing aids. Other extensive data on the population were reported.

Anderson (1967) reported an incidence of 4:1,000 for high-frequency hearing loss among schoolchildren tested in a routine hearing screening program. Children with high-frequency hearing loss exhibited learning failures and behavior difficulties more frequently than did a control group of children. It was concluded that the teacher should always be informed about any child with this defect so that suitable placement in the classroom could be made and appropriate follow-up maintained.

### Special Populations

Webb and others (1966) reviewed 10 major surveys utilizing pure-tone hearing testing of mentally retarded individuals in institutions. Although the incidence of hearing loss appeared to vary according to the chronological and mental age of the subjects, the most striking finding was in variation due to differing criteria for designating what constituted a hearing loss. There was a discussion of the problem of the validity of hearing test procedures on retardates and recommendation for subsequent studies. Lloyd and Reid (1967) reported pure-tone audiometric data on 638 institutionalized mentally retarded individuals ranging from 6-22 years of age. One hundred and thirty-eight had a hearing loss of greater than 15 dB (ASA-1951) at one or more of seven frequencies between 250 and 8,000 Hz. The data were related to previous investigations, and medical and non-medical implications were discussed. Dansinger and Madow (1966) administered the Verbal Auditory Screening for Children (VASC) to 967 retarded children and adults. Of these, 109 (11.3 percent) were untestable. Hearing impairment was reported using several criteria and 12.6 percent of those tested were found to have sufficient hearing impairment to interfere with communication. It was concluded that the VASC is effective in detecting hearing impairment among institutionalized mentally retarded individuals.

### Etiology and Pathology

Hardy (1968) summarized various etiologies of deafness. Sections in her report dealt with genetic determinants, drugs, noise, prenatal viral infections, and postnatal factors. Of particular interest was the information on the rubella epidemic which swept the Eastern seaboard in 1963-64 and spread across the United States, reaching the West Coast in 1964-65. Children deafened by that epidemic are presently seeking admission to educational programs and will be a source of considerable concern in future years since many of them have other disabilities in addition to deafness.



Vernon (1968a) reported on various etiologies involved in deafness and their psychological and educational concomitants. His study presented information on current etiological factors in deafness based on 1,468 children who had applied to a state residential school during the period 1953-64. Tables showing prevalence rates for various etiologies were presented. The leading causes of deafness and their ranges of possible prevalence were found to be heredity (5.3-26.1 percent), prematurity (11.9-17.5 percent), meningitis (7.7-8.7 percent), maternal rubella (8.8-9.5 percent), and Rh factor (3.1-3.7 percent). In 30.4 percent of the cases, cause was unknown, and in 32.3 percent another and/or an etiology of dubious validity was given. Changing etiologies of deafness have resulted in a school-age population today which is about 5 percent postlingually deafened in contrast to past years when it was approximately 40 percent.

### **Audiologic Assessment**

The literature on audiologic assessment was too extensive to permit detailed review. In general, investigations continued the trends in infant assessment, objective audiometry, and differential diagnosis noted by Rosenstein in the February 1966 REVIEW OF EDUCATIONAL RESEARCH.

#### **Infant Assessment**

Ewertsen (1966) reported an investigation of screening audiometry with infants ranging from 6-15 months in age. Two stuffed animals with speakers mounted in their stomachs were positioned on either side of the infant being tested. Shifts in attention during presentations of warbled tones were used to estimate hearing level. A conditioned audiovisual response technique was described by Reddell and Calvert (1967). The test was designed for children between the ages of 1 and 2½ years. Description and validity data obtained with a group of 20 hearing-impaired children were presented.

Barnet and Lodge (1966) described the use of computer-averaged electroencephalographic responses to sound with infants. Results of EEG testing with 22 infants correlated highly with those obtained by clinical evaluation of the infants' behavior in response to auditory stimuli. Downs and Sterritt (1964) studied identification audiometry with neonates using the eye-blink and/or the Moro reaction to auditory stimuli with 117 infants. Responses of the infants to two types of auditory stimuli were judged by two or three observers. High rates of agreement among the observers were reported. Downs and Sterritt (1967) also reported the development of a community program of detection of hearing loss at birth. The fact that the program may produce a numerically low yield was believed to be far outweighed by the benefits to the individual child.

## Objective Audiometry

Davis (1966) compared threshold levels of 79 deaf pupils, aged 4-13 years, on tests of evoked-response audiometry, clinical audiometry, and voluntary audiometry and found good agreement among the procedures. Price and Goldstein (1966) reported the results of using averaged evoked EEA responses to obtain auditory thresholds on 70 children, aged 2 months through 13 years. With the exception of a tendency toward larger amplitudes and longer latencies, waveform of the children's responses was similar to that of adults. Sedation with Nembutal did not appear to affect averaged evoked responses; however, Pentothal often resulted in records difficult to analyze.

## Differential Diagnosis

Owens (1965a) administered the Short Increment Sensitivity Index (SISI) to 27 subjects with normal hearing, 95 patients with cochlear lesion, 15 patients with VIIIth nerve lesion, and three patients of uncertain classification because of discrepancies between the audiologic and the neurologic findings. On the whole, the test appeared highly useful in determining site of lesion. Owens (1965b) also investigated the question whether the SISI and Alternate Binaural Loudness Balance (ABLB) were duplicative in the sense that both may reflect loudness recruitment. From the results it was inferred that the two tests were duplicative. Burke and Creston (1966) tested subjects with sensorineural noise-induced hearing loss with a fixed-frequency Bekesy-type audiometer. Results indicated that this type of technique might be the most fruitful approach to recruitment testing.

## Other Studies

Davis (1965) reported the recommendations of the Committee on Conservation of Hearing of the American Academy of Ophthalmology and Otolaryngology on the division of hearing handicap into classes or grades according to a table included in the article. The committee also strongly recommended that every recorded audiometric reading be identified clearly as either ASA-1951 or ISO. The ISO reference zero applies to air-conduction pure-tone audiometers. No corresponding standard has been developed for bone conduction. The committee recommended that no change be made in speech audiometers in their reference zero levels at the present time. The speech reception thresholds they measure should agree with predictions from ISO pure-tone audiometry at least as well as, and probably much better than, they did with ASA-1951 pure-tone audiometry.

Griffing, Simonton, and Hedgecock (1967) reported a study of 175 children using the Verbal Auditory Screening for Children (VASC). The data indicated that 89.7 percent of the children were accurately classified by this procedure. The authors stated that the use of VASC in preschool testing makes possible early identification of impaired hearing. A pulse-tone

group screening test using conventional audiometric equipment was administered by Hollien and Thompson (1967) to 234 college students in groups of up to 20 students. No known cases of hearing loss were missed by the group technique. High validity was found on a comparison of the results of 152 children tested by individual audiometry.

Elliott (1967) presented a statistical description of routine audiometric and psychometric scores of the 177 children enrolled in the school divisions of a private residential school during the academic year 1964-65. Comparisons of earliest audiograms with the most recent ones showed remarkable stability. When ordinary audiometers were supplemented by separate oscillators and amplifiers with consequent higher power outputs, only five children failed to respond at 1,000 Hz or lower frequencies and 88 percent responded at 2,000 Hz. Uneven performance was observed on subtests of both intelligence and achievement test batteries.

## Psychological Studies

There was a considerable increase during the past five years in the number of studies concerned with psychological aspects of deafness. In addition to individual studies from a number of investigators, there were continuing efforts by Vernon (1968b) in the areas of assessment and etiology; by Furth (1966a, b; 1968) in nonverbal learning and cognition; by Kates, Kates, and Michael (1962) in cognitive processes; and by Rainer and Altshuler (1966) in mental health.

### Assessment

Vernon (1968b) summarized 50 years of research on the intelligence of deaf children and children who were hard of hearing. He discussed various types of tests used and made recommendations concerning them. Montgomery (1966) reported the relationship between tests of residual hearing, intelligence, and achievement and admittance to educational programs. Evans (1966) conducted a comparative study of the *Wechsler Intelligence Scale for Children (WISC)* (Performance) and Raven's *Progressive Matrices*. The tests were administered to 100 children who were deaf or had partial hearing, and three years later 42 children were retested. The *WISC* (Performance) results showed high internal consistency and high retest stability and were distributed within reasonably normal limits. The *Progressive Matrices* test was also a highly reliable measure, but the results were substantially below normal and showed some fluctuation with time. There was a low intercorrelation between the two tests.

Personality testing of the handicapped was discussed by Wachs (1966). Five areas of handicapping conditions were considered: blindness, deafness, speech disorders, motor disorders, and intellectual retardation. A summary of each handicapping condition indicated which tests seemed most promising for diagnosis. Cowen and others (1967) reported a network of three



studies related to the development of a scale to assess attitudes toward deafness.

The performance scales of the Wechsler tests continued to be the most widely used intelligence tests with deaf children. Raven's *Progressive Matrices* seemed to have gained in popularity, apparently on the assumption that it might assess in nonverbal form more abstract types of abilities than are measured by performance tests. Personality assessment was concentrated on the use of projective techniques rather than paper and pencil tests, most of which appear to have limited utility with deaf persons because of their deficient language ability.

### Perceptual and Motor Skills

Suchman (1968) reported on the frequency and educational consequences of visual impairment among a group of more than one hundred deaf children. More than half of the children were found to have some visual impairment which reflected a higher frequency than usually reported for deaf children. Most of the visually impaired children who had correctable vision were receiving no medical treatment. Binnie, Elkind, and Steward (1966) compared hearing-impaired subjects and subjects with normal hearing on visual perceptual ability. Differences were found between subgroups of hearing-impaired subjects based on age at onset and whether they were oral or nonoral subjects.

Tests of visual memory span were administered by Olsson and Furth (1966) to adolescents who were deaf and who had normal hearing and adults equated for age and intelligence. Consistent results at both age levels indicated that the deaf subjects and the subjects with normal hearing performed similarly with nonsense forms, but that the deaf subjects were inferior with digits. The results were interpreted to support the role of previous experience rather than verbal mediation in immediate recall. Hartman and Elliott (1965) found that 24 deaf children performed less well than a matched group of children with normal hearing on a response-alternation memory task which was hypothesized to be facilitated by auditory imagery.

Suchman (1966) studied color-form preference, discrimination accuracy, and learning of deaf children and children with normal hearing. Pairs of deaf subjects and subjects with normal hearing were matched for age, sex, and IQ and were compared (a) on ability to discriminate accurately in the color and the form dimensions and (b) on a successive discrimination learning task. Subjects with normal hearing discriminated among forms more accurately than deaf subjects, and deaf subjects discriminated more accurately among colors than subjects with normal hearing. Subjects with normal hearing learned the form discrimination with fewer errors than deaf subjects. The investigator discussed preference-response characteristics and how preference relates to perceptual discrimination and learning behaviors. Sterritt, Camp, and Lipman (1966) studied nine

hearing-impaired children and nine children with normal hearing using a telegraph key to reproduce temporal patterns created by an above-threshold tone or by a flashing light. Subjects with hearing losses were inferior to the control subjects in reproducing the auditory temporal patterns. They also were inferior in their ability to reproduce the visual temporal patterns, suggesting that the effects of sensory deprivation were not limited to the deprived sensory modality.

Boyd (1967) compared deaf boys and boys with normal hearing on the motoric functions of static equilibrium, locomotor coordination, and psychomotor integration and explored the area of laterality in detail. The deaf subjects were inferior to the control group on static equilibrium and locomotor coordination. Thirty percent of the endogenous deaf group preferred the left hand indicating a possible genetic link between deafness and sinistrality.

### Cognitive Functioning

Considerable research was published in recent years on the nonverbal and cognitive functioning of deaf people. Most of this research has been summarized by Furth (1968) in a review and perspective on the thinking of deaf people.

Furth (1966a, b; 1968) and a group of associates have conducted research for a number of years on the thought processes of deaf people. They have concluded from their research and that of others that language is not necessary in the thinking process. A similar conclusion was reached by Kates and a group of associates (1962) who also conducted a continuous series of investigations in this area. The two groups of investigators tended to disagree on the interpretations of their investigations with regard to the education of deaf children. Furth has contended that (a) language is not a constituent element of logical thinking, (b) prelingually deaf persons are linguistically incompetent in the formal language of the hearing community, (c) logical thinking can be developed in deaf children by non-verbal means and by conventional gesture (sign) language, and (d) formal language of the hearing community should be taught to deaf children after they have acquired proficiency in the conventional gesture (sign) language commonly used by many deaf persons. These contentions were criticized by Blank (1965) who considered that Furth had presented inadequate evidence for stating that thinking could develop without language. She stated that many deaf subjects, possibly including those used by Furth, are well trained in language and thus cannot be considered linguistically deficient. Furth disputed the contentions of Blank regarding the linguistic competence of deaf persons. Kates (1967) summarized the work of his group on the cognitive structures of deaf persons and persons with normal hearing and added comparative data on the cognitive processes of psychotics.

Nunnally and Blanton (1966) compared the word associations of two groups of deaf subjects and a group of subjects with normal hearing. The

results indicated that the deaf gave more associations that could have been learned from visual experience and reading and fewer associations that supposedly are learned from subtle nuances of language. It was suggested that, as a group, words were less meaningful to the deaf subjects than to those with normal hearing. Koplin and others (1967) compared two groups of deaf subjects with a group of subjects having normal hearing on the Kent-Rosanoff word associations test. The results were consistent with the hypothesis that word associations of the deaf are comparable to those of younger subjects with normal hearing. Blake, Ainsworth, and Williams (1967) compared 54 deaf subjects and 54 subjects with normal hearing on the effects of induction and deduction on the attainment of first-order concepts. The subjects with normal hearing responded more adequately than the deaf subjects in recognizing and using the concepts.

## Mental Health

Rainer and his associates continued their work on mental health and psychotherapy with deaf persons. Their most recent book (Rainer and Altshuler, 1966) presented the implications of their research for the development of comprehensive mental health services. Martorano and Oestreicher (1966) described the use of hypnotic techniques with 12 deaf, mentally ill patients. Implications for the further use of hypnosis with deaf persons were discussed. Denmark (1966) reported a study of 28 profoundly deaf patients in mental hospitals. It was observed that the incidence of deaf persons in mental hospitals in Britain exceeded the incidence for the general population and that rediagnosis of the patients showed the initial diagnosis appearing on hospital records to be frequently incorrect.

The past and continuing research on adapting therapeutic procedures for use with deaf persons holds promise for the eventual large-scale provision of mental health services for the deaf.

## Language and Communication

Since the February 1966 REVIEW OF EDUCATIONAL RESEARCH, a number of major works appeared in the literature which reported or summarized research relating to the development of language and communication in the deaf and persons who were hard of hearing. Pickett (1968) edited the proceedings of the Conference on Speech-Analyzing Aids for the Deaf held in June 1967. The publication contained papers by individuals from many countries concerning the use of technological aids in the development of speech, speechreading, and residual hearing. It could well represent the beginning of a major trend in the development of such aids and techniques for use with deaf individuals. Quigley (1966b) prepared an issue of the *Volta Review* which featured the contributions of several



authors summarizing the research literature on language acquisition by deaf children and children with normal hearing. New directions in the study of the language development of deaf children were indicated in a monograph by Rosenstein and MacGinitie (1965). The several articles in the monograph were concerned with the application of psycholinguistic techniques to the study of language development in deaf children. West (1968) edited a report containing translations of Russian work in the field of speech and hearing, including auditory impairment.

The sheer number of articles in these publications precluded any comprehensive discussion of them in this review. Only a few have been selected to indicate what might be major trends for the future in research on language and communication in deaf children and children who are hard of hearing. Five such trends appeared to be developing: (a) the design and construction of devices to transform the acoustic signal of speech into another sensory modality (usually visual but also vibratory in some cases) which would provide an additional feedback to replace or supplement the deficient auditory feedback of the deaf person, (b) transformation of the acoustic signal of speech into a visual signal to serve as a receptive means of communication to replace speechreading or improve the teaching of it, (c) the development of devices to produce a frequency shift in the acoustic signal to utilize whatever residual hearing a deaf child might have, (d) an increased interest in the use of manual communication (fingerspelling and the language of signs) in educating deaf children, and (e) application of recent developments in linguistics and psycholinguistics to the language and communication problems of the deaf child.

## Speech

Liberman and others (1968) discussed the basic problems involved in providing the deaf person with a useful nonacoustic representation of the speech signal. They concluded that visual representation in the form of speech spectrograms had so far met with little success and that the best visual displays of speech for the deaf person might be representations of articulatory muscle contractions involved in speech production. Børrild (1968) and Risberg (1968) described a number of speech-analyzing devices by means of which important information could be extracted from the speech signal and displayed visually. Visual displays consisting of spectrum indicators; fricative indicators; and intonation, rhythm, and nasalization indicators were described. Mártony (1968) compared measurements of voice pitch (fundamental frequency) between children who were severely hard of hearing and children with normal hearing. Using a modified version of Tjernlund's pitch extractor to provide a visual display of voice pitch, he reported encouraging results with four severely hard of hearing subjects who were able to use the device to establish control of pitch and then to transfer this control to kinesthetic (or auditory) feedback alone, without any visual information.

A number of studies used the new Bell Telephone Laboratories' Visible Speech Translator (VST) to discover what features of speech might be taught to deaf children with spectral and amplitude contour displays. Stark, Cullen, and Chase (1968) used this device with eight severely or profoundly deaf subjects who ranged in age from 3-20 years. They obtained improved intelligibility with most subjects, but greatest success was obtained in school-age subjects, and some carry-over into everyday speech for those subjects was observed. House, Goldstein, and Hughes (1968) employed a version of the VST to teach subjects to identify spoken vowel patterns on a spectrographic display. Improved performance by a group of 25 subjects in identifying vowel patterns from spectrographic displays was noted.

Pronovost (1968) described a prototype voice visualizer for speech training and its preliminary evaluation with deaf children. Studies of vowels with 12 children and consonants with 12 other children resulted in articulatory improvement as measured by listener preference. However, among the children with severe hearing loss and poor speech there was little improvement in intelligibility. Phillips and others (1968) used visual displays of voice pitch portrayed as a function of time on a screen to teach intonation to deaf subjects by visual pattern matching. Preliminary analysis indicated that some improvement in intonation patterns was achieved. Kringlebotn (1968) described experiments with visual and vibrotactile aids in teaching speech to deaf children. Experiments with an s-sound indicator giving visual or vibrotactile indication of the s-sound showed that the aid was useful in teaching this sound to deaf children.

Boone (1966) compared 44 deaf children and 44 children with normal hearing on a number of acoustical and physiological variables associated with speech production. He found that the voices of the deaf children as compared to those of the children with normal hearing showed alterations of pitch level, resonance, and phonation duration. Suggestions were offered as to how these aspects of speech might be improved by the teacher.

Cornett (1967) developed a method known as cued speech which might be of value in the teaching of speech as well as speechreading. The method consists of four hand positions and eight hand configurations (or cues) which provide information to the deaf person on those features of speech which are not readily visible. By reading what is visible on a speaker's lips, supplemented by the cues for those features which are not visible, the deaf persons supposedly should be able to readily understand what is being said. Conversely, the cues can be used to instruct the child concerning nonvisible speech features in his own speech which are in need of improvement.

### Speechreading

Research on speechreading continued to be scanty and concentrated on the development of tests and the exploration of factors which might be related to speechreading ability. Donnelly and Marshall (1967) discussed

the development of a multiple-choice test of speechreading with two equivalent forms. Meaningful divisions within the test were derived to assist in homogeneous therapy placement of deaf college students. Butt and Chreist (1968) developed a speechreading test for very young children. Data obtained from 130 hearing-impaired children ranging in age from 3 to 9 years indicated the test was valid and reliable and suitable for children at the prereading level. Stepp (1966) studied the feasibility of using programed instruction in speechreading for self-study by acoustically handicapped children. The results indicated that the 8mm motion pictures as used in the project with 10 hard of hearing and profoundly deaf children were effective in providing speechreading practice for individual study.

Evans (1965) presented speechreading films to 64 congenitally deaf or prelingually deafened children ages 8-16 years and related the scores to various psychological factors. He found that visual recognition, intelligence, and degree of hearing loss were the principal predictive variables, with visual recognition being the best predictor of speechreading potential. Schwartz and Black (1967) studied the effects of sentence structure on speechreading performance. Six kernel sentences and various transformations of them were presented to subjects with normal hearing in the presence of masking noise. Results indicated the transformational rules of generative grammar related systematically to success in speechreading.

A new development was reported by Upton (1968). He described a wearable eyeglass hearing aid which could provide visual cues of distinctive speech features not readily visible in speechreading. An electronic analyzer was used to extract voicing, fricative, and stop information from speech. Miniature lights mounted on eyeglasses worn by the subject were caused to flash in synchronism with the speech to form dynamic light patterns representative of the presence of those speech features. The cues provided aided the subject in speechreading. They also were of value in providing the subject with visual feedback of certain features of his speech output. The similarity between this device and cued speech (Cornett, 1967) will be noted. Cued speech and the eyeglass speechreading aid both attempt to provide visual information of distinctive speech features not readily visible in speechreading.

### **Amplification and Auditory Training**

There was continued research on the effects of amplification for persons with hearing impairment but little corresponding work on auditory training. The use of amplification was concerned with two questions: (a) Is amplification of value to deaf persons with very little residual hearing? and (b) What type of frequency amplification is best? These questions are continuations of the issues raised during the past 20 or 30 years through the educational use of amplification with profoundly deaf children. It would seem that the first question had been adequately answered in the affirmative, yet it continues to appear as a subject of research.



Hirsh (1966) summarized recent developments in the use of hearing aids including stereophonic listening, selection of hearing aids, and greater use of hearing aids. A cautionary note was sounded by Macrae and Farrant (1965) who studied the effects of high-powered hearing aids on the residual hearing of a group of children with sensorineural deafness. It was concluded that the use of such aids tended to cause deterioration of residual hearing. This finding supported earlier work by Kinney (1961). It was questioned by Bellefleur and Van Dyke (1968) who studied the audiometric records of 58 deaf children covering a 10-year period and found no deterioration of residual hearing because of amplification. Barr and Wedenberg (1965) followed 84 children with bilateral "perceptive" losses for a period of up to 15 years. Exogenous hearing impairments (maternal rubella or perinatal accidents) did not show any progressivity despite continuous use of hearing aids. Hearing impairments caused by meningitis all showed progressivity. It was suggested that progression of loss was spontaneous rather than caused by the use of amplification.

Several attempts have been made to shift or transpose high-frequency speech sounds, with the aim of making them more audible and intelligible to persons with residual hearing in the low frequencies. Since most deaf children have some residual hearing in the very low frequencies, these experiments could be of considerable importance for future educational practices. Ling and Druz (1967) summarized the research in this area and reported a study of their own comparing results obtained with the use of a transposing instrument and a conventional speech training aid. Results with the transposing instrument did not prove to be superior to those obtained on the speech training aid. Ling (1968) reported three experiments in transposition. One used a form of partial vocoding, one employed the Johansson (1966) transposing instrument, and the other used both. Results of the studies, which assessed auditory discrimination skills among children with low-frequency residual hearing trained on transposition, were largely negative. In spite of the present lack of positive results with transposition amplification, this appears to be a promising area for future research.

Selective amplification to fit individual hearing losses has been a topic of considerable discussion and some research. The basic question has been whether essentially flat amplification of a broad band of frequencies is as good as selective amplification of particular frequencies to fit individual hearing losses. Reddell and Calvert (1966) reported findings on 24 subjects with high-frequency loss and no previous experience with hearing aids. The subjects' responses with aids custom-adjusted to their particular hearing losses proved to be superior to commercially available hearing aids selected by customary audiological procedures.

Much interest has been generated, but little data reported, concerning the use of amplification and auditory training with infants and very young children. One approach which has resulted in considerable discussion is

that of the Hearing Education through Auditory Research (HEAR) Foundation. Griffiths (1967) described the rationale, test techniques, and therapy procedures of this approach. Forty-two case histories were presented to illustrate improvements in hearing sensitivity that occurred as a result of early amplification and therapy.

Gaeth and Lounsbury (1966) reported on the use of hearing aids by 134 elementary school children who were hard of hearing. Interviews, test results, and measurements of acoustic characteristics of aids indicated that no more than 50 percent of the children were getting adequate hearing by the most lenient standards. Parent counseling and thorough training with the child and his hearing aid were recommended.

Lichtenberg (1966) investigated the effects of auditory training on vowel and consonant discrimination of children who were hard of hearing. She found that intensive auditory training significantly improved discrimination ability in the children as compared with a control group of hard of hearing subjects who did not receive training.

### Manual Communication

During the past 10 years there has been increased interest and research in the use of manual communication in educating deaf children. This interest developed from the research in Russia during the 1950's on the use of neo-oralism (fingerspelling and speech). The Russian research was summarized by Morkovin (1960) who reported that the Russians, in working with deaf children over a period of several years starting at the age of 2, claimed to have succeeded in developing vocabularies of several thousand words by the time the children were 6 years old. The Russians also claimed that the use of fingerspelling fostered the development of speech and speechreading and that the deaf children in their experiments eventually were able to abandon fingerspelling and rely solely on speech and speechreading for communication.

Quigley and Frisina (1961) published a study concerning the effects of institutionalization on deaf children. The study contained some data relating to the influence of manual communication. They found that deaf children with deaf parents had poorer speech intelligibility but had significantly larger vocabularies and better educational achievement than deaf children who had parents with normal hearing. Stuckless and Birch (1966) compared two matched groups of deaf children, one group having been exposed to manual communication early in life because of having deaf parents and the other group not having had early exposure to manual communication. The findings showed that the group with early exposure to manual communication was superior to the comparison group in reading, written language, and speechreading and was similar to the comparison group in speech intelligibility and psychosocial development.

Quigley (1966a) reported preliminary findings of a five-year investigation of the use of fingerspelling with deaf children in eight residential

schools. The children taught by the combined use of fingerspelling and speech were found to be superior to the comparison groups (without fingerspelling) on most variables in which meaningful language was involved. There were no systematic and significant differences between the fingerspelling and nonfingerspelling groups on speech intelligibility and speechreading ability.

Meadow (1968) compared the development of self-concept in a group of deaf students having deaf parents and a group of deaf students having parents with normal hearing. She found that the group with deaf parents had better development of self-concept and better psychosocial adjustment than did the group with parents who had normal hearing. The group of students with deaf parents also had significantly higher academic achievement and reading achievement and were superior in written language, fingerspelling, and the use of the language of signs. The two groups did not differ on speechreading and speech ability.

Another group of studies of manual communication was concerned with analysis of the linguistic code of the language of signs and with development of methods for teaching the sign language and fingerspelling. Stokoe (1960) performed a linguistic analysis of the language of signs and attempted to describe its cherology (analogous to phonology in spoken language) and its syntactical structure. A further development of this work was a dictionary by Stokoe, Casterline, and Croneberg (1965) which presented a method for transcribing the language of signs and included an extensive lexicon of signs. Bornstein (1965) developed a filmed program for teaching the manual alphabet and reported on its use with a group of subjects.

### Language Development

Recent research on the language development of the deaf child has begun to use the newer linguistic and psycholinguistic techniques. Rosenstein and MacGinitie (1965) edited a monograph containing studies by several authors using these techniques.

Quigley and Moores (1967) used "cloze" techniques with deaf subjects by deleting every fifth word from selected reading passages and requiring the subjects to provide the missing words. The deaf subjects were significantly inferior to a control group with normal hearing in their response to the task on semantic and syntactic measures. Odom, Blanton, and Nunnally (1967) also employed "cloze" procedures with deaf subjects and found them to be inferior to a control group with normal hearing, with the inferiority being greater on syntactic than on semantic measures.

Brannon and Murray (1966) obtained spoken language samples of 30 sentences from 30 subjects with hearing impaired and 30 subjects with normal hearing. Fifteen of the hearing-impaired subjects were classified as hard of hearing and 15 as deaf. The deaf subjects had significantly lower total word output than did the other two groups. Both the deaf subjects



and the subjects who were hard of hearing had significantly lower structural accuracy than the subjects with normal hearing. Brannon (1968) used the linguistic classification system of Jones, Goodman, and Wepman (1963) to analyze the spoken language of three groups of subjects: those with normal hearing, those who were hard of hearing, and those who were deaf. It was concluded that a significant hearing impairment reduced productivity of both tokens and types of words. Moderate impairment lowered the use of adverbs, pronouns, and auxiliaries; a profound impairment reduced nearly all classes.

Stuckless (1966a) collected written language samples from a stratified sample of students in schools for the deaf throughout the United States. Various objective measures were computed and correlated with judgments of the written samples by three teacher-judges. The three descriptive variables which collectively had the highest correlation with teacher judgments were grammatic correctness ratio, type-token ratio, and composition length. Multiple regression equations incorporating these three measures were developed.

Recent research on language development indicated there had been a shift from the analysis of language samples by traditional grammatical methods to application of psycholinguistic techniques in the study of the dynamics of language development in the deaf child and the relation of these to the dynamics of language development in the child with normal hearing.

### **Educational Studies**

The extensive literature on the education of deaf children contained very little objective research. The few studies that were published were scattered among a number of educational areas, and there were no attempts to examine any important areas, such as programed learning or operant conditioning, in a systematic manner.

Quigley, Jenné, and Phillips (1968) reported a survey study of 653 deaf individuals and individuals who were hard of hearing who attended 326 accredited institutions of higher education in 45 states. Factors related to success in college attendance were discussed, and recommendations were made concerning special provisions that might aid increased numbers of hearing-impaired persons to benefit from higher education facilities for the general population. Salem (1967) reported on the abilities and progress of seven graduates of a residential school for the deaf who were in attendance in a regular college. Vaughn (1968) reported the results of a study designed to provide supportive services for deaf students and students who were hard of hearing in attendance at a college and a vocational school for the general population. Extensive data were presented on the educational progress and occupational placement of the students.

Elliott and Armbruster (1967) discussed some possible effects of delay of early treatment of deafness. Questionnaire responses were obtained from parents of children enrolled in a school for the deaf which indicated that those children who had learning problems in addition to severe hearing impairment had received treatment services at a later age than the other children in the school. The data were interpreted as suggesting the hypothesis that sensory deprivation in hearing-impaired children resulting from delayed educational procedures as well as delayed sound amplification may produce an additional learning handicap which is only partly reversible by later therapeutic procedures.

Furth (1966a) reported data adapted from a comprehensive survey of the reading ability of deaf children conducted by Wrightstone, Aronow, and Moskowitz (1963). The data showed that between the ages of 11 and 16 years the mean grade equivalent scores of the deaf students advanced only from grade 2.7 to grade 3.5, or less than one grade level in five years. The data also revealed that by the age of 16 years only about 12 percent of the students in the national sample were scoring reading grade levels higher than grade 4.9.

## Social and Vocational Studies

Greatly increased emphasis on the rehabilitation of deaf people has resulted in an increased number of research projects on the social and vocational aspects of deafness. Much of this research was summarized by Stuckless (1966b) in the proceedings of a research conference. Recent research projects were concentrated in the areas of assessment, social adjustment, and community services.

### Assessment

Pimentel (1967) described the use of the Testing, Orientation, and Work Evaluation in Rehabilitation (TOWER) system in the vocational evaluation of deaf persons. Applications and results of the system, composed of more than one hundred specific job sampling tests in 13 broad job families, were given. The use of projective techniques in personality evaluation of deaf adults was discussed by Brenner and Thompson (1967). Limitations of personality instruments for assessment of deaf children and adults were studied by Rosen (1967). His research with the *Minnesota Multiphasic Personality Inventory* led to the conclusion that procedures dependent upon language responses are not suitable for personality assessment of deaf persons. Rosen (1968) also studied the preferences of deaf college students regarding the hearing status of counselors. One in five of the subjects made a clear choice for a deaf counselor. The remaining 80 percent preferred a counselor with normal hearing, a deaf counselor if necessary for adequate communication, or were impartial. Phillips and Berg (1967) presented information on the use of the SRA Verbal and

Non-Verbal Forms with deaf college students over a period of approximately ten years. Hurwitz (1968) developed a classification system of vocational counseling goals and recommended methodologies as a means of objectifying counselor knowledge of practice and providing a common frame of reference in dealing with deaf persons of working age.

### Social Adjustment

Rodda (1966) reported an investigation of 23 hearing-impaired students who left school in England. Information was provided on general background, occupational expectations and provisions, social adaptability and mobility, and personality organization. The results of interviews with the children, their parents, and their teachers showed that the children had a different pattern of adjustment than nonhandicapped children. Craig (1965) studied the effects of deafness and of institutionalization on the development of the self-concept. She found that the accuracy of the self-concept of the deaf child was hampered by his language deficit, regardless of his residence in an institution or his home. The self-acceptance of the institutionalized deaf group was significantly higher than for a deaf non-institutionalized group and a nondeaf, noninstitutionalized group.

### Community Services

A number of research and demonstration projects have explored the utilization of community services for deaf people. The October 1967 issue of the *Journal of Rehabilitation of the Deaf* is devoted to articles describing several such projects. The research on the development of community services seems to represent a major trend in the future provision of services for deaf persons at the community level.

### Bibliography

- ANDERSON, URSULA M. "The Incidence and Significance of High-Frequency Deafness in Children." *American Journal of Diseases of Children* 113: 560-65; May 1967.
- BARNET, ANN B., and LODGE, ANN. "Diagnosis of Deafness in Infants with the Use of Computer-Averaged Electroencephalographic Responses to Sound." *Journal of Pediatrics* 69: 753-58; November 1966.
- BARR, BENGT, and WEDENBERG, ERIK. "Prognosis of Perceptive Hearing Loss in Children with Respect to Genesis and Use of Hearing Aid." *Acta Oto-Laryngologica* 59: 462-74; May 1965.
- BELLEFLÉUR, PHILIP A., and VAN DYKE, ROBERT C. "The Effects of High Gain Amplification on Children in a Residential School for the Deaf." *Journal of Speech and Hearing Research* 11: 343-47; June 1968.
- BINNIE, CARL; ELKIND, DAVID; and STEWARD, JOSEPH. "A Comparison of Visual Perceptual Ability of Acoustically-Impaired and Hearing Children." *International Audiology* 5: 238-41; No. 2, 1966.
- BLAKE, KATHRYN A.; AINSWORTH, STANLEY H.; and WILLIAMS, CHARLOTTE L. "Effects of Induction and Deduction on Deaf and Hearing Individuals' Attainment of First-Order Concepts." *American Annals of the Deaf* 112: 606-13; September 1967.
- BLANK, MARION. "Use of the Deaf in Language Studies: A Reply to Furth." *Psychological Bulletin* 63: 442-44; June 1965.



- BOONE, DANIEL R. "Modification of the Voices of Deaf Children." *Volta Review* 68: 686-90; November 1966.
- BORNSTEIN, HARRY. *Reading the Manual Alphabet*. Washington, D.C.: Gallaudet College Press, 1965. 78 pp.
- BØRRILD, KNUD. "Experience with the Design and Use of Technical Aids for the Training of Deaf and Hard of Hearing Children." *American Annals of the Deaf* 113: 168-77; March 1968.
- BOYD, JOHN. "Comparison of Motor Behavior in Deaf and Hearing Boys." *American Annals of the Deaf* 112: 598-605; September 1967.
- BRANNON, JOHN B., JR. "Linguistic Word Classes in the Spoken Language of Normal, Hard-of-Hearing, and Deaf Children." *Journal of Speech and Hearing Research* 11: 279-87; June 1968.
- BRANNON, JOHN B., JR., and MURRAY, THOMAS. "The Spoken Syntax of Normal, Hard-of-Hearing, and Deaf Children." *Journal of Speech and Hearing Research* 9: 604-10; December 1966.
- BRENNER, LEON, and THOMPSON, RICHARD E. "The Use of Projective Techniques in Personality Evaluation of Deaf Adults." *Journal of Rehabilitation of the Deaf* 1: 17-30; July 1967.
- BURKE, K. S., and CRESTON, J. E. "Recruitment in Noise-Induced Hearing Loss." *Acta Oto-Laryngologica* 62: 351-61; October-November 1966.
- BUTT, DOLORES, and CHEIST, FRED. "A Speechreading Test for Young Children." *Volta Review* 70: 225-39; April 1968.
- CORNETT, R. ORIN. "Cued Speech." *American Annals of the Deaf* 112: 3-13; January 1967.
- COWEN, EMORY L., and OTHERS. "Development and Evaluation of an Attitudes to Deafness Scale." *Journal of Personality and Social Psychology* 6: 183-91; June 1967.
- CRAIG, HELEN B. "A Sociometric Investigation of the Self-Concept of the Deaf Child." *American Annals of the Deaf* 110: 456-82; May 1965.
- DANSINGER, STUART, and MADOW, ARNOLD A. "Verbal Auditory Screening with the Mentally Retarded." *American Journal of Mental Deficiency* 71: 387-92; November 1966.
- DAVIS, HALLOWELL. "Guide for the Classification and Evaluation of Hearing Handicap in Relation to the International Audiometric Zero." *Transactions of the American Academy of Ophthalmology and Otolaryngology* 69: 740-51; July-August 1965.
- DAVIS, HALLOWELL. "Validation of Evoked-Response Audiometry (ERA) in Deaf Children." *International Audiology* 5: 77-81; No. 2, 1966.
- DENMARK, JOHN C. "Mental Illness and Early Profound Deafness." *British Journal of Medical Psychology* 39: 117-24; Part 2, 1966.
- DONNELLY, KENNETH G., and MARSHALL, WILLIAM J. A. "Development of a Multiple-Choice Test of Lipreading." *Journal of Speech and Hearing Research* 10: 565-69; September 1967.
- DOWNES, MARION P., and STERRITT, GRAHAM M. "Identification Audiometry for Neonates: A Preliminary Report." *Journal of Auditory Research* 4: 69-80; April 1964.
- DOWNES, MARION P., and STERRITT, GRAHAM M. "A Guide to Newborn and Infant Hearing Screening Programs." *Archives of Otolaryngology* 85: 15-22; January 1967.
- ELLIOTT, LOIS L. "Descriptive Analysis of Audiometric and Psychometric Scores of Students at a School for the Deaf." *Journal of Speech and Hearing Research* 10: 21-40; March 1967.
- ELLIOTT, LOIS L., and ARMBRUSTER, VIRGINIA B. "Some Possible Effects of the Delay of Early Treatment of Deafness." *Journal of Speech and Hearing Research* 10: 209-24; June 1967.
- EVANS, LONEL. "Psychological Factors Related to Lipreading." *Teacher of the Deaf* 63: 131-36; May 1965.
- EVANS, LONEL. "A Comparative Study of the Wechsler Intelligence Scale for Children (Performance) and Raven's Progressive Matrices with Deaf Children." *Teacher of the Deaf* 64: 76-82; March 1966.
- EWERTSEN, H. W. "Teddy-Bear Screening Audiometry for Babies." *Acta Oto-Laryngologica* 61: 279-80; March 1966.
- FURTH, HANS G. "A Comparison of Reading Test Norms of Deaf and Hearing Children." *American Annals of the Deaf* 111: 461-62; March 1966. (a)
- FURTH, HANS G. *Thinking Without Language: Psychological Implications of Deafness*. New York: Free Press, 1966. 236 pp. (b)

- FURTH, HANS. "A Review and Perspective on Thinking of Deaf People." *Cognitive Studies*. Vol. 1. Seattle, Wash.: Special Child Publications. (In press)
- GAETH, JOHN H., and LOUNSBURY, EVAN. "Hearing Aids and Children in Elementary Schools." *Journal of Speech and Hearing Disorders* 31: 283-89; August 1966.
- GENTILE, AUGUSTINE; SCHEIN, JEROME D.; and HAASE, KENNETH. *Characteristics of Persons with Impaired Hearing*. U.S. Department of Health, Education, and Welfare, Public Health Service, Publication No. 1000, Series 10, No. 35. Washington, D.C.: Government Printing Office, 1967. 64 pp.
- GRIFFING, TERRY; SIMONTON, KINSEY; and HEDGECOCK, LEROY. "Verbal Auditory Screening for Preschool Children." *Transactions of the American Academy of Ophthalmology and Otolaryngology* 71: 105-10; January-February 1967.
- GRIFFITHS, CIWA. *Conquering Childhood Deafness*. New York: Exposition Press, 1967. 112 pp.
- HARDY, JANET B. "The Whole Child: A Plea for a Global Approach to the Child with Auditory Problems." *Education of the Deaf: The Challenge and the Charge*. Washington, D.C.: Government Printing Office, 1968. pp. 13-34.
- HARTMAN, JACQUELINE S., and ELLIOTT, LOIS L. "Performance of Deaf and Hearing Children on a Short Term Memory Task." *Psychonomic Science* 3: 573-74; December 1965.
- HIRSH, IRA. "Recent Developments in Hearing-Aid Use." *International Audiology* 5: 393-98; No. 3, 1966.
- HOLLIEN, HARRY, and THOMPSON, CARL L. "A Group Screening Test of Hearing." *Journal of Auditory Research* 7: 85-92; January 1967.
- HOUSE, A. S.; GOLDSTEIN, D. P.; and HUGHES, G. W. "Perception of Visual Transforms of Speech Stimuli: Learning Simple Syllables." *American Annals of the Deaf* 113: 215-21; March 1968.
- HURWITZ, SIDNEY N. "A Treatment Nosology for Vocational Counseling of the Deaf." *Journal of Rehabilitation of the Deaf* 2: 42-53; April 1968.
- JOHANSSON, BERTIL. "The Use of the Transposer for the Management of the Deaf Child." *Journal of International Audiology* 5: 362-71; No. 3, 1966.
- JONES, LYLE V.; GOODMAN, MORRIS F.; and WEPMAN, JOSEPH M. "The Classification of Parts of Speech for the Characterization of Aphasia." *Language and Speech* 6: 94-107; April-June 1963.
- KATES, SOLIS. *Cognitive Structures in Deaf, Hearing and Psychotic Individuals*. Northampton, Mass.: Clarke School for the Deaf, 1967. 145 pp.
- KATES, SOLIS L.; KATES, WILLIAM W.; and MICHAEL, JAMES. *Cognitive Processes in Deaf and Hearing Adolescents and Adults*. Psychological Monographs: General and Applied, Vol. 76, No. 32. Washington, D.C.: American Psychological Association, 1962. 34 pp.
- KINNEY, CHARLES E. "The Further Destruction of Partially Deafened Children's Hearing by the Use of Powerful Hearing Aids." *Annals of Otolaryngology and Laryngology* 70: 828-35; September 1961.
- KOPLIN, JAMES H., and OTHERS. "Work Association Test Performance of Deaf Subjects." *Journal of Speech and Hearing Research* 10: 126-32; March 1967.
- KRINGLEBOTN, M. "Experiments with Some Visual and Vibrotactile Aids for the Deaf." *American Annals of the Deaf* 113: 311-17; March 1968.
- LIBERMAN, A. M., and OTHERS. "Why Are Speech Spectrograms Hard to Read?" *American Annals of the Deaf* 113: 127-33; March 1968.
- LICHTENBERG, FRANCES S. "A Comparison of Children's Ability To Make Speech Sound Discriminations." *Volta Review* 68: 426-34; June 1966.
- LING, DANIEL. "Three Experiments on Frequency Transposition." *American Annals of the Deaf* 113: 283-94; March 1968.
- LING, DANIEL, and DRUZ, WALTER S. "Transposition of High Frequency Sounds by Partial Vocoding of the Speech Spectrum: Its Use by Deaf Children." *Journal of Auditory Research* 7: 133-44; April 1967.
- LLOYD, LYLE L., and REID, MICHAEL J. "The Incidence of Hearing Impairment in an Institutionalized Mentally Retarded Population." *American Journal of Mental Deficiency* 71: 746-63; March 1967.
- MACRAE, J. H., and FARRANT, R. H. "The Effect of Hearing-Aid Use on the Residual Hearing of Children with Sensorineural Deafness." *Annals of Otolaryngology and Laryngology* 74: 409-19; June 1965.

- MÁRTONY, JÁNOS. "On the Correction of the Voice Pitch Level for Severely Hard of Hearing Subjects." *American Annals of the Deaf* 113: 195-202; March 1968.
- MARTORANO, JOSEPH T., and OESTREICHER, CHARLES. "Hypnosis of the Deaf Mentally III: A Clinical Study." *American Journal of Psychiatry* 123: 605-606; November 1966.
- MEADOW, KATHRYN. "Early Communication in Relation to the Deaf Child's Intellectual, Social and Communicative Functioning." *American Annals of the Deaf* 113: 29-41; January 1968.
- MONTGOMERY, GEORGE W. "Differences in the Interrelationships of Non-Verbal Intelligence Test Scores, Educational Attainments and Residual Hearing in Selected and Unselected Populations." *Teacher of the Deaf* 64: 29-33; January 1966.
- MORKOVIN, BORIS V. "Experiment in Teaching Deaf Preschool Children in the Soviet Union." *Volta Review* 62: 260-68; June 1960.
- NUNNALLY, JUM C., and BLANTON, RICHARD L. "Patterns of Word Association in the Deaf." *Psychological Reports* 18: 87-92; February 1966.
- ODOM, PENELOPE B.; BLANTON, RICHARD L.; and NUNNALLY, JUM C. "Some 'Cloze' Technique Studies of Language Capability in the Deaf." *Journal of Speech and Hearing Research* 10: 816-27; December 1967.
- OLSSON, JAMES E., and FURTH, HANS G. "Visual Memory-Span in the Deaf." *American Journal of Psychology* 79: 480-84; September 1966.
- OWENS, ELMER. "The SISI Test and VIIIth Nerve Versus Cochlear Involvement." *Journal of Speech and Hearing Disorders* 30: 252-62; August 1965. (a)
- OWENS, ELMER. "The SISI Test and Recruitment of Loudness by Alternate Binaural Loudness Balance." *Journal of Speech and Hearing Disorders* 30: 263-68; August 1965. (b)
- PHILLIPS, NATHAN D., and OTHERS. "Teaching of Intonation to the Deaf by Visual Pattern Matching." *American Annals of the Deaf* 113: 239-46; March 1968.
- PHILLIPS, ROBERT, and BERG, THOMAS. "Use of the SRA Verbal and Non-Verbal Forms at Gallaudet College." *Journal of Rehabilitation of the Deaf* 1: 59-62; October 1967.
- PICKETT, J. M. "Sound Patterns of Speech: An Introductory Sketch." *American Annals of the Deaf* 113: 120-26; March 1968.
- PIMENTEL, ALBERT T. "The TOWER System as a Vocational Test for the Deaf Client." *Journal of Rehabilitation of the Deaf* 1: 26-31; April 1967.
- PRICE, LLOYD L., and GOLDSTEIN, ROBERT. "Averaged Evoked Responses for Measuring Auditory Sensitivity in Children." *Journal of Speech and Hearing Disorders* 31: 248-56; August 1966.
- PRONOVOST, WILBERT, and OTHERS. "The Voice Visualizer." *American Annals of the Deaf* 113: 230-38; March 1968.
- QUIGLEY, STEPHEN P. "The Problems of Communication of Deaf Children." *What Is Special Education?* Middlesex, England: Association for Special Education, 1966. Section V, pp. 263-70. (a)
- QUIGLEY, STEPHEN P., issue editor. "Language Acquisition." *Volta Review* 68: 1-131; January, 1966. (b)
- QUIGLEY, STEPHEN P., and FRISINA, D. ROBERT. *Institutional and Psycho-Educational Development of Deaf Children*. Research Monograph. Washington, D.C.: Council for Exceptional Children, 1961. 49 pp.
- QUIGLEY, STEPHEN P., and MOORES, DONALD F. "'Cloze' Procedures in Assessment of Language Skills of Deaf Persons." *Proceedings of the International Conference on Oral Education of the Deaf*. Vol. II. Washington, D.C.: Alexander Graham Bell Association for the Deaf, 1967, pp. 1363-82.
- QUIGLEY, STEPHEN P.; JENNÉ, WILLIAM C.; and PHILLIPS, SONDRÁ B. *Deaf Students in Colleges and Universities*. Washington, D.C.: Alexander Graham Bell Association for the Deaf, 1968. 196 pp.
- RAINER, JOHN D., and ALTSHULER, KENNETH Z. *Comprehensive Mental Health Services for the Deaf*. New York: Department of Medical Genetics, New York State Psychiatric Institute, Columbia University, 1966. 191 pp.
- REDELL, RAYFORD C., and CALVERT, DONALD R. "Selecting a Hearing Aid by Interpreting Audiologic Data." *Journal of Auditory Research* 6: 445-52; October 1966.
- REDELL, RAYFORD, and CALVERT, DONALD R. "Conditioned Audio-Visual Response Audiometry." *Voice* 16: 52-57; No. 2, 1967.
- RISBERG, ARNE. "Visual Aids for Speech Correction." *American Annals of the Deaf* 113: 178-94; March 1968.



- RODDA, M. "Social Adjustment of Hearing-Impaired Adolescents." *Volta Review* 68: 279-83; April 1966.
- ROSEN, ALBERT. "Limitations of Personality Inventories for Assessment of Deaf Children and Adults as Illustrated by Research with the *Minnesota Multiphasic Personality Inventory*." *Journal of Rehabilitation of the Deaf* 1: 47-52; July 1967.
- ROSEN, ALBERT. "Deaf College Students' Preferences Regarding the Hearing Status of Counselors." *Journal of Rehabilitation of the Deaf* 1: 20-27; January 1968.
- ROSENSTEIN, JOSEPH. "The Deaf and the Hard of Hearing." *Review of Educational Research* 36: 176-98; January 1966.
- ROSENSTEIN, JOSEPH, and MACGINITIE, WALTER, editors. *Research Studies on the Psycholinguistic Behavior of Deaf Children*. Research Monographs, Series B, No. B02. Washington, D.C.: Council for Exceptional Children, 1965. 40 pp.
- SALEM, JAMES M. "Deaf Students in a 'Hearing College.'" *Volta Review* 69: 36-40; January 1967.
- SCHWARTZ, JOYCE, and BLACK, JOHN W. "Some Effects of Sentence Structure on Speechreading." *Central States Speech Journal* 18: 86-90; May 1967.
- STARK, RACHEL E.; CULLEN, JOHN K.; and CHASE, RICHARD A. "Preliminary Work with the New Bell Telephone Visible Speech Translator." *American Annals of the Deaf* 113: 205-14; March 1968.
- STAPP, ROBERT E. "A Speechreading Laboratory for Deaf Children." *Volta Review* 68: 408-15; June 1966.
- STERRITT, GRAHAM M.; CAMP, BONNIE WEBB; and LIPMAN, BEVERLY S. "Effects of Early Auditory Deprivation upon Auditory and Visual Information Processing." *Perceptual and Motor Skills* 23: 123-30; August 1966.
- STOKOE, WILLIAM C., JR. *Sign Language Structure*. Buffalo: Department of Anthropology and Linguistics, University of Buffalo, 1960. 78 pp.
- STOKOE, WILLIAM, JR.; CASTERLINE, DOROTHY C.; and CRONEBERG, CARL G. *A Dictionary of American Sign Language on Linguistic Principles*. Washington, D.C.: Gallaudet College Press, 1965. 346 pp.
- STUCKLESS, E. ROSS. "Objective Evaluation of Original Written Language of Deaf Students." *Volta Review* 68: 679-85; November 1966. (a)
- STUCKLESS, E. ROSS, editor. *Research on Behavioral Aspects of Deafness*. Vocational Rehabilitation Administration, U.S. Department of Health, Education, and Welfare. Washington, D.C.: Government Printing Office, 1966. 121 pp. (b)
- STUCKLESS, E. ROSS, and BIRCH, JACK W. "The Influence of Early Manual Communication on the Linguistic Development of Deaf Children." *American Annals of the Deaf* 111: 452-60; March 1966.
- SUCHMAN, ROSSLYN G. "Color-form Preference, Discriminative Accuracy and Learning of Deaf and Hearing Children." *Child Development* 37: 439-51; June 1966.
- SUCHMAN, ROSSLYN G. "Visual Impairment Among Deaf Children—Frequency and Educational Consequences." *Volta Review* 70: 31-37; January 1968.
- U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE. *Education of the Deaf: The Challenge and the Charge*. Washington, D.C.: Government Printing Office, 1968. 157 pp.
- UPTON, HUBERT W. "Wearable Eyeglass Speechreading Aid." *American Annals of the Deaf* 113: 222-29; March 1968.
- VAUGHN, GWENYTH. "Hearing-Impaired Students Benefit from Supportive Programs." *Volta Review* 70: 14-23, 58; January 1968.
- VERNON, MCCAY. "Fifty Years of Research on the Intelligence of Deaf and Hard-of-Hearing Children: A Review of Literature and Discussion of Implications." *Journal of Rehabilitation of the Deaf* 1: 1-12; January 1968. (b)
- VERNON, MCCAY. "Current Etiological Factors in Deafness." *American Annals of the Deaf* 113: 106-15; March 1968. (a)
- WACHS, THEODORE D. "Personality Testing of the Handicapped: A Review." *Journal of Projective Techniques and Personality Assessment* 30: 339-55; August 1966.
- WEBB, CLARENCE, and OTHERS. "Incidence of Hearing Loss in Institutionalized Mental Retardates." *American Journal of Mental Deficiency* 70: 563-68; January 1966.
- WEST, ROBERT, editor. *Russian Translations on Speech and Hearing*. Washington, D.C.: American Speech and Hearing Association, 1968. No. 3, 432 pp.
- WRIGHTSTONE, J. WAYNE; ARONOW, MIRIAM S.; and MOSKOWITZ, SUE. "Developing Reading Test Norms for Deaf Children." *American Annals of the Deaf* 108: 311-16; May 1963.

## Index to Volume XXXIX, No. 1

Page citations, though made to single pages, often indicate the beginning of a chapter, section, or running discussion dealing with a topic.

- Aphasia: 45; diagnosis and therapy for, 47; language, cognition, and perception studies, 45; lesion site and speech and language behavior, 46
- Articulation disorders: 38; description and diagnosis of, 39; motor, sensory, and perceptual factors, 38; therapy for, 40
- Audiologic assessment: 105; differential diagnosis, 106; infant assessment, 105; objective audiometry, 106; other studies, 106
- Behavior disorders: 83; characteristics, 88; prevalence, 84; prospective and suggested research, 97; teacher behavior, training, and role, 95; therapeutic educational provisions, 91
- Behavior disorders, characteristics of: 88; school achievement, 88; social and psychological correlates, 89
- Behavior disorders, prevalence of: 84; epidemiology in the public schools, 85; persistence of emotional disturbance, 86; screening and predictive measures, 84
- Creativity of the gifted, developing of: 30
- Creativity research: 26
- Deaf and hard of hearing: 103; audiologic assessment, 105; educational studies, 117; etiology and pathology, 104; incidence, 103; language and communication, 110; psychological studies, 107; social and vocational studies, 118
- Deaf and hard of hearing, educational studies: 117
- Deaf and hard of hearing, etiology and pathology: 104
- Deaf and hard of hearing, incidence of: 103; general population, 103; special populations, 104
- Delayed speech and language: 44; diagnosis and therapy, 44; perceptual factors associated with, 44
- Gifted: 25; creativity research, 26; developing creativity, 30; differences among, 33; educational practices, 32; identifying, 27; social influences and giftedness, 25; teaching of, 28
- Gifted, differences among: 33
- Gifted, educational practices of: 32
- Gifted, identifying the: 27
- Language and communication: 110; amplification and auditory training, 113; language development, 116; manual communication, 115; speech, 111; speechreading, 112
- Learning disabilities: 71; research methodology, 80; research in transition, 71; correlates of, 72; special education methods, 78
- Learning disabilities, correlates of: 72; age-related, 74; neurochemical, 77; psychological, 76; prediction, 72
- Mentally retarded: 5; classification, 5; education and rehabilitation, 8; psychological processes, 11
- Mentally retarded, classification of: 5; individual intelligence tests, 5; nature of mental retardation, 5; nonstandardized instruments and direct measurement, 7; short-form intelligence scales, 6; sociometric and personality inventories, 7
- Mentally retarded, education and rehabilitation of: 8; behavior modification, 10; curricular adaptations, 8; placement, 8; postschool adjustment, 10
- Mental retardation, psychological processes of: 11; language and communication, 16; learning, 11; memory, 16; personality, 17; social adjustment, 17;
- Multiple handicapped: 76
- Prospective and suggested research: 97
- Psychological studies: 107; assessment, 107; cognitive functioning, 109; mental health 110; perceptual and motor skills, 108
- Research methodology: 80
- Research in transition: 71

- Special education methods: 78
- Speech handicapped: 38; aphasia, 45; articulation disorders, 38; delayed speech and language, 44; stuttering, 40; voice disorders, 42
- Social and vocational studies: 118; assessment, 118; community services, 119; social adjustment, 119
- Social influences and giftedness: 25
- Stuttering: 40; adaptation and stuttering, 41; characteristics of stutterers, 41; linguistics variables and stuttering, 41; modification of, 42
- Teacher behavior, training, and role: 95
- Teaching of the gifted: 28
- Therapeutic educational provisions: 91; approaches to behavioral change and management, 94; effects of special class placement, 93; models for learning and behavior changes, 92; prevention and school consultation, 94
- Visually handicapped: 52; education, 59; general aspects, 52; multiple handicapped, 66; psychological aspects, 55; technological research, 54
- Visually handicapped, education of: 59; listening, 59; mobility, 65; sight utilization, 64; tactile reading, 61
- Visually handicapped, general aspects of: 52
- Visually handicapped, psychological aspects of: 55; abstract functioning and concept development, 57; social acceptance, 55; tests and appraisal, 55
- Visually handicapped, technological research on: 54
- Voice disorders: 42; analysis of, 43; identification and evaluation of, 43; treatment of, 43





# ABSTRACTS OF PAPERS 1969

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

1969 ANNUAL MEETING

Los Angeles, California, February 5-8, 1969

All papers read at the AERA 1969 Annual Meeting will be summarized in 300-word abstracts and bound into a pocket-size paper-covered book which may be purchased for \$2.50 in the registration area at the Biltmore Hotel.

After the Annual Meeting—from February 10, 1969—orders for copies of the ABSTRACTS will be filled at AERA headquarters in Washington, D.C., at \$3.00. Single copy orders must be accompanied by check or money order. Supplies will be limited, so mail the coupon as soon as you can after February 10.

---

## Publication Sales

American Educational Research Association

1126 Sixteenth Street, N.W.

Washington, D.C. 20036

Please send \_\_\_\_\_ copies of ABSTRACTS OF PAPERS 1969, American Educational Research Association Annual Meeting @ \$3.00 per copy. Enclosed is a check/money order for \$ \_\_\_\_\_.

---

NAME and TITLE

---

STREET ADDRESS

---

CITY

STATE

ZIP CODE





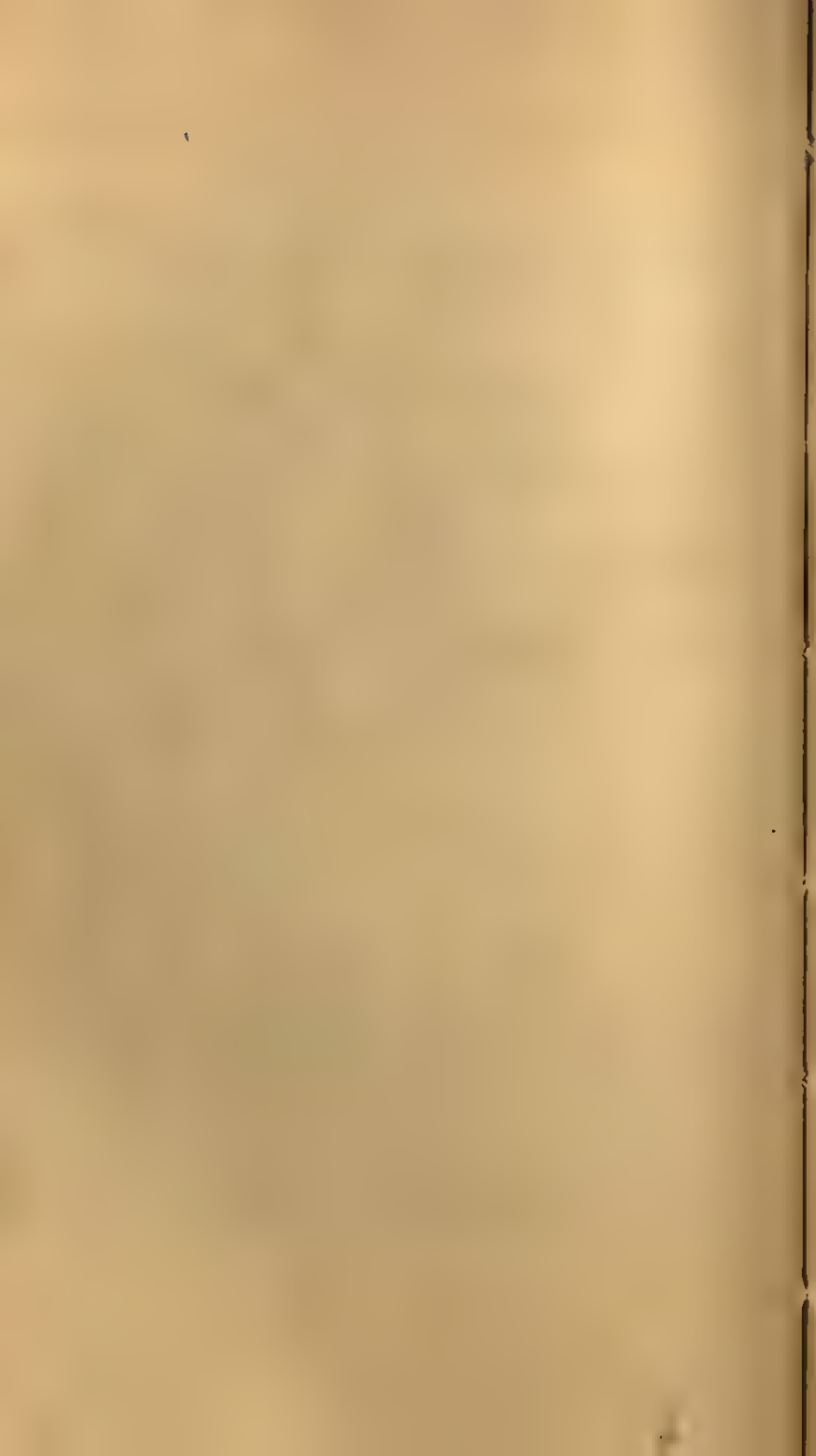
## FOREWORD

Research in counseling, guidance and personnel services as it is designed and carried out has much to be modest about. Most studies fail to ask or explore questions that in any way alter theoretical rationales, programs or practices. The authors who contributed chapters to this issue were asked to select studies carefully, to review them thoroughly, and to comment upon them critically and in detail. Authors were encouraged to focus on ways of making research more relevant to theory and practice, in part by identifying good examples from the research literature. Where none were found, authors have suggested needed studies and new lines of inquiry. Authors have also identified the types of studies and the problem areas in which further inquiry seems futile. A primary purpose of this issue was that of making research more relevant to theory and practice.

New chapters have been added on the topics of systems approaches in guidance, changes or outcomes of counseling, and relevancy in counseling research. Some chapters in the April 1966 issue of the *Review* have been retitled and the chapter on the appraisal function was dropped. A new procedure was initiated with this issue. The authors were aided by Consultants who were selected for their recognized competence in the topic area. Another innovation was the contribution of the Counseling and Personnel Services Information Center (ERIC) at the University of Michigan. The ERIC Center under the direction of Garry Walz assisted several of the authors in searching, compiling and reproducing documents not otherwise available. The Cooperative Community Educational Resources Center of Boulder, Colorado also provided bibliographic search support.

The efforts of several persons deserve special acknowledgement. The individual authors merit praise for their hard work, especially for their good-natured acceptance of the "editorial assistance" they received. The contribution of the Consultants is much appreciated. The editorial assistance of Jack Hamilton and William Stilwell deserves special comment, as does the encouragement, consideration and patience of Gene Glass, the General Editor. The services as consultants to the General Editor of Drs. Albert Roark, William Sease and Theodore Volsky of the University of Colorado are gratefully acknowledged. Finally, the excellent work of David Shaw who assisted with all phases of this Issue is much appreciated.

Carl E. Thoresen  
*Issue Editor*



# 1: THEORETICAL FOUNDATIONS OF GUIDANCE

Martin R. Katz\*

*Educational Testing Service*

In the previous triennial review on this topic, Kehas (1966) identified two divergent views of the relationship between guidance and theory. One is an "atheoretical orientation, crediting guidance with no independent substance and, hence, no need for its own theory." The other is a "substantive orientation," which posits that guidance is not merely a "derivative of more basic disciplines," but warrants distinctive theories of its own. Kehas characterized the work emanating from the latter orientation as mainly, "discussions of theory" with some presentations of theoretical constructs, but no "full-blown theoretical systems of guidance." Theory has matured during the last three years, but is not yet full blown.

## *Metatheory*

As a helping profession, guidance requires a theory of *intervention*. The foundation for a theory of intervention should include some understanding of the phenomena with which the intervention is concerned—for example, what is the nature of career decisions and how are they made in the absence of a planned professional intervention? For an answer to

---

\*Dr. Carroll Miller, Northern Illinois University, and Dr. Chris Kehas, Claremont Graduate School, served as consultants to Dr. Katz on the preparation of this chapter. Ruth Ekstrom assisted in compiling publications for this review.



this question, the guidance profession looks to research in career development. Studies in the "natural history" of career development, however, while relevant for guidance theory, are not sufficient for it—just as theories of learning are not sufficient for a theory of instruction. Some of the topics to be covered in theoretical systems of guidance might include definitions, purposes, contents, methods and processes, outcomes, participants and settings (Katz, 1967). Of these, career development theory would pertain mainly to the content and outcomes of guidance. Career development theory may help in recognizing stages of development—in describing and explaining what is happening. The distinctive element in guidance theory goes beyond discovery and explanation of what is happening and leads into a rationale for what we (who?) want (why?) to make (how?) happen (did it?). "Theory" has often been used as an antithesis to practice; but guidance requires development of a theory for practice.

Sprinthall and Tiedeman (1966) suggested that a comprehensive theory of guidance might emerge from a "consolidation of counseling theory and career development," but they were not explicit about how the consolidation might take place. If guidance involves intervention in people's lives, there would seem to be an obligation to be as explicit as possible about the nature of the intervention. Theory is, after all, an attempt to make implicit beliefs explicit.

The convention is to support such beliefs with evidence. This leads to conclusiveness in theories. Another convention is to support beliefs with logical connections to other beliefs. This leads to inclusiveness. One may provide depth; the other, breadth, to borrow Oppenheim's (1957) nomenclature.

Theories, like persons, go through stages of development. Examining the state of theory at a given time may serve as a crucial safeguard against premature hardening of practice. Breadth without depth can lead to shallow eclecticism that often comes from knowing too little about too much, but narrow depth can lead to the faddism that comes from knowing too much about too little.

For that matter, a theory that is relatively broad and deep, if it becomes rigidified, is no guarantee against quackery. When practitioners lose perspective on the state of their theory, they may "know" too much about too much. Bakan (1968) recreated what may be regarded as a full-blown theory of guidance—phrenology. According to his analysis, phrenological theory had the flavor of modern science. It had breadth: concepts of anatomy and physiology were related to behavior, and the cortical localization of functions permitted measurement, diagnosis, prescription and training in accordance with "individual differences." Phrenology also acquired considerable depth by testing numerous propositions experimentally. It produced many successful outcomes. Bakan credited phrenology with bringing about humane treatment of the insane—"exercising" a weakened faculty rather than "literally beating the devil out of

the afflicted." Other significant and distinctive interventions based on phrenological theory included educating blind deaf-mutes previously regarded as ineducable, advocacy of "learning by doing" and exposing infants to a multitude of sensations and stimuli. The theory of the skull eventually "failed," and vocational guidance as practiced by phrenologists became a pet subject for castigation in the literature of guidance through the 1930's.

Clearly, a "full-blown theory of guidance" is not without limitations. Any theory of intervention can be dangerous, but so can atheoretical intervention. In a beautifully written book, Kaplan (1964, p. 296) points up the uncertain (but irresistible) rewards of theory:

... a theory ... reaches out beyond itself. ... The possibility of failure is intrinsic to its effort, and the aspiration to truth may be paid for by facing the ever-present risk of error, as the danger of death is the price exacted of life. ... error, to be sure, is of man's making[:]; so, in the same sense, is truth—"nothing ventured, nothing gained" is also sound epistemology.

It is appropriate to note, in this connection, Bakan's claim that phrenology paved the way for the "American psychology" of James, Dewey and Angell. Perhaps this is the most useful function of all theories—paving the way for new ones. Theory may be superior to brute empiricism primarily because it helps men to learn more from their mistakes.

By virtue of metatheory, then, a theory may be useful even if it lacks veridicality. It must be remembered that a theory is not reality. It is the way we construe the world. It makes our premises, as well as our conclusions, explicit. It may deal, therefore, with what we believe we know and also with what we know we believe.

Still, it must be noted that some theories of action contain greater risks than others. The first rule of medicine is to do no harm. (Perhaps also, the heavy reliance placed on the counselor's judgment accounts for some of the resistance to "behavioral" counseling.) It is not enough that a system be explicit. The most rigid dogmas can be highly explicit. But it is the rigidity, not the explicitness that should be feared. As Muller (1957, p. 70) said, "... pretensions to absolute certainty ... are the ultimate source of corruption, the reason why the best becomes the worst and crusaders for heaven make a hell on earth." Hence, the more explicit a system is *and* the more open to extension, modification, or even discard, the less harmful it may be. Of course, *harmlessness* is not a sufficient standard for a theory of intervention. What other standards have been invoked?

### *Standards and Judgments*

Efforts have been made to establish standards and pass judgment on the state of theory. Carkhuff (1966) and Carkhuff, Alexik and Anderson

(1967) gave counseling theory and "vocational choice theory" low marks, according to the rules of an inductive-deductive model specified in the latter article. The efforts reviewed in both areas—130 publications on counseling and 5 programs of research on "vocational choice"—were sternly denied the accolade of "theory." In the latter article, the authors complained of the lack of connection between low-level generalizations based on data and high-level constructs. In a judgment on the judges, however, the authors' discussion of scientific method was characterized, in passing, as "jejune" by Roe (1968). The present reviewer merely wonders, how can an inductive-deductive model set standards for judging at what stage of development theory *should be* in a given field at a given time? If theoretical efforts have not come up to these standards, does this mean that there are no theories?

In this connection, it is appropriate to note Bakan's (1967) argument that hypothesis testing is no royal road to psychological understanding. Oppenheim (1957), a distinguished philosopher of science, defined a number of bipolar characteristics of theory—such as broad or deep, theoretical or observational, typifying or individualizing, specialistic or universalistic, nomothetic or idiographic, more or less balanced, strong or weak—in terms of two basic dimensions: *extensity* and *strength*. However, it should be emphasized that these characteristics are descriptive, not evaluative.

To write of the theoretical foundations of guidance may seem pretentious to those who feel that research in guidance has much to be modest about. However, critics are perhaps more pretentious when they establish rigid ground rules to stipulate which sets of ideas may be dubbed "theory." "Scientism" is not science. Although hypothesis-testing should not be scorned, neither should the element of play with ideas in theory. A search for symmetry, aesthetic pleasure, and simplicity ("Seek simplicity and distrust it," said Whitehead) has made important contributions to the advancement of theory in the sciences as well as in the arts.

As Kaplan (1964) said, "Every scientist is first of all a poet, giving to airy nothing a local habitation and a name." But the interventionist is not so much a poet as a dramatist. He is concerned less with nouns than with verbs, and he is particularly concerned with the adverbs *why* and *how*.

### *Terms, Definitions and Purposes*

As soon as men give something a name, they express or imply some theory about it. Sometimes nomenclature changes with changes in the popularity of a theory (cf. "occupational choice" and "career development"), but sometimes names are the potsherds of theories now deceased or moribund. Self-consciousness of counselors about their professional status may increase the mortality rates of terms. Only a little more than a decade ago, Super (1955) described the transition from "Vocational Guidance" to



"Counseling Psychology." More recently, Tiedeman (1967), in his presidential address to the Division of Counseling Psychology, APA, urged that the profession exhume "guidance" and bury "counseling." Counseling psychology, he said, had focused on techniques of counseling, and had not been sufficiently concerned with defining goals "within the framework of education."

In school settings, *counseling* (a process) may be regarded as a major element in *guidance* (a program), which in turn may be considered a component of *pupil personnel services* (an administrative designation about which no more will be said in this chapter). By synecdoche, however, "counseling" has often stood for "guidance" in the literature. It seems defensible, therefore, in much of this chapter to treat the terms as overlapping, if not interchangeable. Indeed, practitioners are called "guidance counselors" and "counselors" quite indiscriminately. However, it must be recognized that "counseling" is sometimes perceived to be more extensive than "guidance": it is occasionally found in the company of such modifiers as vocational, educational, academic, developmental, behavioral, adjustment, personal, social, ethical and financial. These same adjectives are almost as frequently associated with "guidance," however, suggesting that the most pressing business is not to pit "counseling" and "guidance" against each other. Of greater concern is a distinctive definition of the territory that guidance and counseling jointly occupy in education.

Historically, the field of guidance has been inundated by successive floods of educational purposes. Perhaps the high-water mark of non-definition was reached with such statements as "Guidance is education, and education is guidance," and "All counselors are teachers, and all teachers are counselors." In the sediment left by these waters have sprouted "special" activities—all designated counseling or guidance—to improve reading, health, citizenship, morality, manners, study habits, etc. This state of affairs led Katz (1967) to propound the bull, "If counseling is what a secondary school counselor does, it includes a host of activities besides counseling." The suggestion was made that "guidance" and "counseling" be reserved for professional intervention in the choices open to a student. Society has pre-determined what some decisions should be. It unequivocally favors good reading, good health and other such "universals" of education. Although counselors may sometimes supplement the work of teachers in transmitting these values, their distinctive concern is not with the "universals" but with the "alternatives"—toward which the culture tends to be more permissive. (The terms "universals" and "alternatives" are borrowed from Smith, Stanley and Shores, 1957). The "alternatives" represent choices between competing values:

... if the role of education is to transmit the culture, an important role of guidance is to help the individual come to terms with the culture—that is, the choices he makes will indicate how



he sees himself in the culture. But first he must see the culture in himself. Thus, his first question should be, where have my values come from? Then he will be better prepared to ask, where are they taking me? (Katz, 1966, p. 4)

In a similar vein, Tiedeman and Field (1965) delineated distinctive but overlapping roles for teacher and counselor in developing students' "purposeful action," which they have stated to be the purpose of guidance. Tiedeman and Field (1965, p. 257) listed three prerequisites for the development of goal-seeking behavior: the student must "1) know of goals and their bases as favored by others, 2) experience the expectation that he will learn how to evolve goal-directed activity of his own accord, and 3) continually subject his wishes and expected responsibilities for purposeful action to critical examination." They maintained that the teacher should work primarily in the realm of the first two of these conditions, and the counselor work in the latter two. Thus, teaching and guidance would complement each other in helping the student to cope with discontinuities in his career.

Fitzgerald (1965) reported that over 90% of the total membership of the American School Counselor Association (ASCA) approved a Statement of Policy that defined counseling as "an accepting, non-evaluative relationship in which [a pupil] is helped to better understand himself, the environment he perceives, and the relationship between these," its purpose being "that most pupils will enhance and enrich their personal development and self-fulfillment by means of making intelligent decisions."

### *Behavioral Goals*

In a series of stimulating and provocative papers, Krumboltz (1965b, 1966a, 1966b) appeared to reject clients' self-understanding and self-acceptance as satisfactory goals of counseling. Instead, Krumboltz proposed as goals those specific and observable behavior changes that are desired by each client and are compatible with each counselor's values. He and his associates have given considerable attention to methods for accomplishing behavior changes; these methods will be considered later in this chapter. There has been, perhaps, less clarity in Krumboltz's deemphasis of self-understanding in the formulation, definition and acceptance of goals. It is difficult to see how the client can adequately designate desirable changes in behavior unless he first perceives his present state and "proceives" the new state to which he aspires. Granted, answering these questions of self-understanding—Who am I? What do I want to become? What is the nature of the discrepancy between these two states?—is not the only goal of guidance. There remain the planning and the accomplishment of the change.

It may also be granted that, as Krumboltz (1966b) has said, behavioral changes can take place (for example, through reinforcement) *without self-*

understanding. But it is difficult, then, to demonstrate that these changes were in accordance with the counselee's goals. Does *specific* behavior change alone suffice? One of the generally accepted aims of guidance is to work itself out of a job—to help students grow increasingly able to solve their own problems, independently and self-sufficiently. Some insight into how the change took place would appear to be desirable for future self-direction. Furthermore, in deprecating self-understanding, does not Krumboltz tend to limit perception of the counselee's goals? The Rogerian would ask, "What keeps the counselee from engaging in appropriate goal-directed behavior? If he wants to get better marks, doesn't he know, as well as the counselor, that he must—let us say—spend more time studying?" What may be overlooked is an implicit value conflict: "I want to get better marks" may leave unsaid "but I don't want to spend more time studying." Krumboltz does not explain how the unspoken clause gets into the counselee's statement of goals or problem, without some attention to self-understanding.

Guidance is not concerned just with the goal that the counselee says he wants at the moment—at report card time, for example. It is concerned with *examined* goals. It tries to stretch the counselee's perceptions without losing sight of present wants, to include a systematic exploration and examination of relevant values which will result in some adjudication between competing or conflicting values. Self-understanding, then, may be not just an instrumental element in changing behavior; it may survive as a higher-order goal of guidance. It is appropriate to note Shoben's (1965) suggestion that problem-solving is less the aim of counseling than its raw material. The chief aim of counseling, in Shoben's view, is "developmental experience."

Krumboltz (1966b) also seems to have slighted the bases on which the individual counselor determines that a counselee's specific goal is "acceptable." Has he avoided the question by simply delegating the responsibility to each counselor without specifying the social, cultural, ethical and individual components of the standards to be used by counselors? To what extent does the counselor's stand depend on his affiliations, for example? To what extent does his stand depend on his own values? If each counselor advertises which kinds of goals he considers compatible with his own values, the counselee may simply shop around until he finds a counselor who agrees with him. This is not likely to challenge the counselee to explore and examine his own goals systematically. Krumboltz's emphasis on problem-solving and behavior-changing as aims of counseling may be better suited for a model of instruction in which the goals are given (the "universals" again) than for a model of guidance in which the goals are to be chosen (the "alternatives"). In the former model, it has been suggested that society (rather than the individual being counseled) is the counselor's real client. If so, the theory should indicate how the counselor helps society to define its goals.

## Existentialism

Theoretical writing in guidance is like a wind-harp activated by ideas that are in the air at a given time, such as behaviorism, existentialism, statistical decision theory, McLuhanacy, etc. This reviewer cannot attempt to cover all of these trends, but feels obligated to acknowledge the existence, at least, of existentialism. Having done so, he must admit failing to answer the question of its distinctive contributions to guidance theory. Buckley's (1966) summary seemed to be more readable than most. Unfortunately, the material that the reviewer was able—after a not inconsiderable struggle—to understand did not seem new, or unique to existentialism. (Unlike Moliere's character who was delighted to find that he had been speaking prose all his life, the present reviewer was somewhat dismayed to recognize that some of his previous writings might be labeled existentialist.) For example, one wonders whether he needed the existentialists to sound what Buckley called this "new note": "... man is not entirely determined . . . he is never more human . . . than when he is making a decision . . ." Among the startling "changes or modifications" such a point of view may have on guidance were listed restoration of "freedom of thought and action" to the counselee and "openness" and "acceptance" on the part of the counselor. (Confronted with these "new" notions, one is tempted to quote Horace: "There were kings before Agamemnon.") Buckley listed seven objectives of existential counseling, lapsing for the most part into the fuzzy lingo that seems to afflict most "existential" writings on guidance. Three of the aims (which this reader thought he would have understood if they had been classified together as one aim) are "to live with problems . . . without undue . . . feelings of frustration, anger or anxiety," "to live with human anxiety . . . to accept limitations . . . as an integral part of being human," and "to accept suffering itself, when inescapable. . . ."

These aims make a nice contrast with the behaviorist emphasis on problems *to be solved*. Social modeling and reinforcement are to the behaviorist like the lever to Archimedes: "Give me a place to stand, and I can raise the world."

Tiedeman (1967) seemed to maintain a foot in each camp by distinguishing between "predicaments to be tolerated" and "problems to be solved." He held that the counselor should not only expect but "cultivate" counselee's anxiety about making choices, that he should help them not only understand their predicament but appreciate it. Tiedeman invoked the importance of engaging the counselee in exercising "the full range of his humanness." An interesting counter-argument against this Aristotelean standard for defining purposes can be found in a paper by Diederich (1951) in which the author went on to propose an ethical basis for educational objectives. Although Diederich did not consider aims of guidance as such, many of the arguments against various bases for educational objectives will prove challenging to those who try to define goals for guidance.



## *Career Development Theory, the Content of Guidance*

It has already been suggested that career development theory contributes primarily to the content rather than to the purposes or methods of guidance. It is not surprising, then, that Osipow's (1968) review of career development theories, although more tolerant than that by Carkhuff, Alexik and Anderson (1967) found few implications for theory, aside from Career Pattern Study, which Super has linked conceptually to guidance. It is somewhat surprising that Osipow noted no similar connection between Tiedeman's work in career development and the computer-based action program for guidance that Tiedeman has purportedly derived from it (Information System for Vocational Decisions, 1967). The work of the Minnesota Work Adjustment Project (for a succinct summary of theory, see Dawis, Lofquist and Weiss, 1968) has also begun to forge some direct links to guidance. In general, however, what career development theory can contribute to guidance is the identification of stages, dimensions, measures and norms. These provide bases for classification, prediction and evaluation that are necessary and useful, despite historical misuses of descriptives as imperatives. (This misuse is analogous to the instructional misconception of mean scores or average grade equivalents as standards that all students should be brought up to.)

Especially noteworthy contributions to measurement of vocational maturity have been made by Crites (1965) and Gribbons and Lohnes (1968). Gribbons and Lohnes found that Readiness for Vocational Planning scores obtained in grade 8 predicted criteria of vocational coping and competence when the students were two years out of high school. This finding can be taken to indicate that certain types of intervention in career decision-making at grade 8 may have long-range as well as short-range effects.

In a manifestation of what may be called neo-trait-and-factor theory, Cooley and Lohnes (1966) contributed to measurement for guidance by using test space (involving the large number of Project TALENT tests) to classify occupations and predict occupational "choice." These classifications and predictions might serve as part of a comprehensive information system for guidance.

Katz (1968) attempted to derive from trends in career development an understanding of what he took to be the *content* of vocational guidance—matters relevant to the choices that society requires or permits. (For example, is the choice between work and non-work now to be classified with the "alternatives" rather than with the "universals"?) These choices were translated by Katz into the outline of a "curriculum" for guidance. Hansen (1967) described more specific curriculum activities that are in keeping with current vocational development theory. O'Hara (1968) derived from vocational development theory the stages of "vocational learning" that might order the use of occupational information in guidance.



Hershenson (1968) indicated some of the implications of his life-stage vocational development system for guidance. Although these publications might all be termed relevant to the "content" of guidance, none dealt specifically with counseling methods.

### *Processes and Methods*

Krumboltz (1965a, p. 226) posed the following question(s):

For clients desiring help on each type of problem of concern to the counselor

What techniques and procedures, When used by what kind of counselors,

With which type of clients, For how long, And in what sequence

Will produce which types of behavior change?

Responses to this litany have been forthcoming in profusion from the reports of a series of experiments undertaken by Krumboltz and his associates. Social modeling and operant conditioning have been the common techniques. Krumboltz and Schroeder (1965) found that reinforcement produced significantly more information-seeking behavior outside the interview for female students in their sample, but not for males, while model-reinforcement counseling produced converse results (the models were male). Thoresen and Krumboltz (1967) examined in some detail the relationships between interview responses and external behavior; they called attention to the complex interactions that remained to be studied among such variables as personality, characteristics of counselors, histories and expectancies of students, etc. Krumboltz, Varenhorst and Thoresen (1967) found that use of female social models did increase the information-seeking behavior of females; but differences in attentiveness and prestige of the model counselor did not affect the criterion behavior. Thoresen, Krumboltz and Varenhorst (1967) attempted to clarify the complex interaction effects of sex of counselors, models and students. Krumboltz (1968) reviewed specific application of behavioral counseling to groups. These seem to be only the beginning, since many additional experiments are foreshadowed by the questions already raised, such as what characteristics of certain counselors make them potent reinforcers for certain students.

Social modeling presumably suggests appropriate behavior so the counselor will not have to wait so long for it to appear before reinforcing it. At the same time, the modeling itself can reinforce the desired behavior vicariously. Of course, similar dynamics may be involved in simply telling students what to do. How well, one wonders, would the experimental groups have compared in information-seeking behavior with control groups that the same counselors had simply *told* to go out and get some information? As Carroll (1968) indicated, most human learning comes "from being told."

Truax (1966) linked behaviorism to Rogerian principles; he found some evidence that therapists who are high in the qualities espoused by Rogers (genuineness, empathy, warmth) may serve as more potent reinforcers than therapists rated low in these qualities. One would still expect to find differences between what the behaviorists and the Rogerians might purport or tend to reinforce in a given case.

Hummel (1966) described "ego-counseling" as a "complex of preferred counselor attitudes and strategies to be implemented flexibly and with respect for the counselee's ultimate freedom. . . ." In method, "ego-counseling" seems to be Rogerian; its definition of "particular sectors" seems close to the behaviorists' concern with accomplishing specific counselee goals or solving specific problems: its "broad sectors" seem to refer to such matters as defining values in connection with decision-making.

### *Decisional Processes*

In the sections on purposes and contents of guidance, this reviewer has neglected much of the literature on guidance as decision-making, largely because most of the basic statements were made before the period covered here. Some developments in method, however, seem worth noting.

Clarke, Gelatt and Levine (1965) and Gelatt (1967) emphasized the role of local research in developing the *informational* and *predictive* components of decision-making; they described the procedures used in the Palo Alto Schools. Gelatt and Clarke (1967) dwelt at some length on the use of subjective probabilities. Thoresen and Mehrens (1967) also considered the question of the influences that different methods of presenting objective probabilities would have on subjective probabilities. Marshall (1967) illustrated the use of the Bayesian approach. Boocock (1967) described role-playing in simulated career decision situations, as embodied in a "life career game." Halpern (1967) devised and tried out the Case Development Questionnaire, a group-administered paper-and-pencil technique for observing students' information-seeking behavior in a simulated situation, which may be modified for instructional use. Katz (1966), who has emphasized the crucial importance of individual values in decision-making, outlined a model that combines a value system, an information system and a prediction system. The role of guidance in the student's examination of values need not be merely maieutic, but can lead him to "try on for size" values that have previously been unfamiliar. Also emphasizing the psychological importance of value systems, Rokeach (1968) described an approach to *changing* students' values—through confrontation of the student with his inconsistencies—that this reviewer believes will be useful as well in a student's explorations of values. The Information System for Vocational Decisions (1967) described a computer-based system for instructing students in career decision-making concepts and for helping them to convert facts and data into relevant and useful information. It emphasized, also, development of the student's "sense of agency."

It is clear that these various decisional approaches have in common a concern with the student's understanding. Cognition—not only about courses of action but about self—seems central in all of them. Whether these decisional approaches contribute to a theory of guidance or not, they seem to require the student's engagement in developing a theory about himself. They imply that theories about self, like theories about guidance, are perhaps most productive when they are both made explicit and kept open. This combination of explicit and open theory seems best to promote learning from experience (which Kaplan, 1964, contrasts with learning *by* experience). Until youth has been exposed to fairly extensive and intensive experiences (real or simulated), hardening of theories about self may be premature.

### Conclusion

Guidance, still struggling through a youthful identity crisis, may also be better off at this time continuing to seek a full-blown theory than finding one.

### Bibliography

- Bakan, David. *On Method: Toward a Reconstruction of Psychological Investigation*. San Francisco: Jossey Bass, 1967.
- Bakan, David. "Is Phrenology Foolish?" *Psychology Today* 1: 44-51; May 1968.
- Boocock, Sarane S. "The Life Career Game." *Personnel and Guidance Journal* 46: 328-334; 1967.
- Buckley, Frank M. "An Approach to Existential Psychology." *Guidance in American Education III: Needs and Influencing Forces*. (Edited by Edward Landy and Arthur M. Kroll). Cambridge: Harvard Graduate School of Education, 82-94; 1966.
- Carkhuff, Robert R. "Counseling Research, Theory and Practice—1965." *Journal of Counseling Psychology* 13: 467-80; 1966.
- Carkhuff, Robert R.; Alexik, Mae; and Anderson, Susan. "Do We Have a Theory of Vocational Choice?" *Personnel and Guidance Journal* 46: 335-45; 1967.
- Carroll, John B. "On Learning From Being Told." *Educational Psychologist* 5: 1, 5-11; March 1968.
- Clarke, Robert; Gelatt, H. B.; and Levine, Louis. "A Decision-Making Paradigm for Local Guidance Research." *Personnel and Guidance Journal* 44: 40-51; 1965.
- Cooley, William W. and Lohnes, Paul R. "Implications for Guidance." *Project Talent One-Year Follow-up Studies*. (Edited by William W. Cooley and J. C. Flanagan.) Cooperative Research Project No. 2333. Pittsburgh: School of Education, Univ. of Pittsburgh, 1966.
- Crites, John O. *Measurement of Vocational Maturity in Adolescence*. Psychological Monographs: General and Applied, Vol. 79. Washington, D. C.: Amer. Psychol. Assoc., 1965.
- Davis, Rene V.; Lofquist, Lloyd H.; and Weiss, David J. *Minnesota Studies in Vocational Rehabilitation: xxiii. A Theory of Work Adjustment (A Revision)*. Minneapolis: Univ. of Minn., 1968.
- Diederich, Paul B. "An Ethical Basis for Educational Objectives." *School Review* 59: 78-86; 1951.
- Fitzgerald, Paul W. "The Professional Role of School Counselors." *Counseling, A Growing Profession*. (Edited by John W. Loughary, Robert O. Stripling, and Paul W. Fitzgerald.) Washington, D.C.: Amer. Personnel and Guidance Assoc., 1965. Pp. 31-41.
- Gelatt, H. B. "Information and Decision Theories Applied to College Choice and Planning." *Preparing School Counselors in Educational Guidance*. New York: College Entrance Examination Board, 1967. Pp. 101-114.



- Gelatt, H. B. and Clarke, Robert B. "The Role of Subjective Probabilities in the Decision Process." *Journal of Counseling Psychology* 14: 332-41; 1967.
- Gribbons, Warren D. and Lohnes, Paul R. *Emerging Careers*. New York: Teachers College Press, 1968.
- Halpern, Gerald. "Assessment of Decision Processes." *Proceedings, 75th Annual Convention, American Psychological Association*. Washington, D. C.: The Assoc., 1967. Pp. 361-62.
- Hansen, Lorraine S. "Theory Into Practice: A Practitioner Looks at Career Guidance in the School Curriculum." *Vocational Guidance Quarterly* 16: 97-103; 1967.
- Hershenson, David B. "Life-Stage Vocational Development System." *Journal of Counseling Psychology* 15: 23-30; 1968.
- Hummel, Raymond C. *An Evaluation of a Model for Guidance Counseling with Bright Academic Underachievers*. Cooperative Research Project No. 5-0633. Pittsburgh: Univ. of Pittsburgh, 1966.
- Information System for Vocational Decisions, *Annual Report 1966-1967*. Cambridge, Mass.: Harvard Graduate School of Education, 1967.
- Kaplan, Abraham. *The Conduct of Inquiry*. San Francisco: Chandler, 1964.
- Katz, Martin. "A Model of Guidance for Career Decision-Making." *Vocational Guidance Quarterly* 15: 2-10; 1966.
- Katz, Martin. *Counseling—Secondary School*. Educational Testing Service Research Memorandum 67-7. Princeton: The Service, 1967.
- Katz, Martin. *The Name and Nature of Vocational Guidance*. Educational Testing Service Research Memorandum 68-1. Princeton: The Service, 1968.
- Kehas, Chris D. "Theoretical Formulations and Related Research." *Review of Educational Research* 36: 207-218; 1966.
- Krumboltz, John D. "Agenda for Counseling." *Journal of Counseling Psychology* 12: 226; 1965a.
- Krumboltz, John D. "Behavioral Counseling: Rationale and Research." *Personnel and Guidance Journal* 44: 383-387; 1965b.
- Krumboltz, John D. "Behavioral Goals for Counseling." *Journal of Counseling Psychology* 13: 153-159; 1966a.
- Krumboltz, John D. *Stating the Goals of Counseling*. California Counseling and Guidance Association, Monograph No. 1. 1966b.
- Krumboltz, John D. "A Behavioral Approach to Group Counseling and Therapy." *Journal of Research and Development in Education* 1: 3-18; 1968.
- Krumboltz, John D. and Schroeder, Wade W. "Promoting Career Planning Through Reinforcement." *Personnel and Guidance Journal* 44: 19-26; 1965.
- Krumboltz, John D.; Varenhorst, Barbara B.; and Thoresen, Carl E. "Nonverbal Factors in the Effectiveness of Models in Counseling." *Journal of Counseling Psychology* 14: 412-418; 1967.
- Marshall, Jon C. "Bayesian Decision." *Journal of Counseling Psychology* 14: 342-345; 1967.
- Muller, Herbert J. *The Uses of the Past*. New York: Oxford Univ. Press, 1957.
- O'Hara, Robert P. "A Theoretical Foundation for the Use of Occupational Information in Guidance." *Personnel and Guidance Journal* 46: 636-640; 1968.
- Oppenheim, Paul. "Dimensions of Knowledge." *Revue Internationale de Philosophie* 40: 1-41; 1957.
- Osipow, Samuel H. *Theories of Career Development*. New York: Appleton-Century-Crofts, 1968.
- Roe, Anne. Letter in *Personnel and Guidance Journal* 46: 1011-1012; 1968.
- Rokeach, Milton. "A Theory of Organization and Change within Value-Attitude Systems." *Journal of Social Issues* 24: 13-33; 1968.
- Shoben, Edward J., Jr. "The Counseling Experience as Personal Development." *Personnel and Guidance Journal* 44: 224-30; 1965.
- Smith, B. Othanel; Stanley, William O.; and Shores, J. Harlan. *Fundamentals of Curriculum Development*. (Revised edition). New York: Harcourt, 1957.
- Sprinthall, Norman A. and Tiedeman, David V. "Guidance and the Pupil." *Year Book of the National Society for the Study of Education*. Sixty-fifth Yearbook. Chicago: The Society, 1966. Part 2, Pp. 59-83.
- Super, Donald E. "Transition: From Vocational Guidance to Counseling Psychology." *Journal of Counseling Psychology* 2: 3-8; 1955.
- Thoresen, Carl E. and Krumboltz, John D. "Relationship of Counselor Reinforcement



- of Selected Responses to External Behavior." *Journal of Counseling Psychology* 14: 140-144; 1967.
- Thoresen, Carl E. and Mehrens, William A. "Decision Theory and Vocational Counseling: Important Concepts and Questions." *Personnel and Guidance Journal* 46: 165-172; 1967.
- Thoresen, Carl E.; Krumboltz, John D.; and Varenhorst, Barbara. "Sex of Counselors and Models: Effect on Client Career Exploration." *Journal of Counseling Psychology* 14: 503-508; 1967.
- Tiedeman, David V. "Predicament, Problem, and Psychology: The Case for Paradox in Life and Counseling Psychology." *Journal of Counseling Psychology* 14: 1-8; 1967.
- Tiedeman, David V. and Field, Frank L. "From a Technology of Guidance in Schools to the Profession of Guidance-in-Society: A Challenge to Democratic Government." *Counselor Development in American Society*. (Edited by John F. McGowan.) Conference Recommendations from Invitational Conference on Government-University Relations in the Professional Preparation and Employment of Counselors, June 2-3, 1965. Washington, D. C. Pp. 249-275.
- Truax, Charles B. "Some Implications of Behavior Theory for Psychotherapy." *Journal of Counseling Psychology* 13: 160-170; 1966.

## AUTHOR

- Katz, Martin R. Address: Educational Testing Service at Princeton, New Jersey Title: Senior Research Psychologist Age: 53 Degrees: B.A., Yale; M.A., Stanford; Ed.D., Harvard Specialization: Guidance, Career Decision-Making.

## 2: SCHOOL GUIDANCE PROGRAMS

H. B. Gelatt\*

*Palo Alto Unified School District*

In this chapter selected articles and publications clearly related to research are discussed and commentary is provided on recent developments in guidance research at the elementary and secondary school levels. In preparing the chapter a variety of sources in addition to the major periodical publications were searched for the period from July 1965 to July 1968. Research evaluating school guidance programs and the role of research as an integral part of such programs are emphasized in this chapter. Since the topic of almost every chapter in this issue is a part, in some way, of school guidance programs, it becomes increasingly difficult and perhaps irrelevant to continue to consider guidance programs separately.

In the last *Review of Educational Research* issue devoted to guidance, Kehas (1966) described guidance as a basic construct which embraces counseling, school psychology, school social work and other activities. He defined related research as that which flows from or leads to particular theoretical formulations on guidance. Theoretical formulations would include those efforts to develop systematic conceptual frameworks for purposes, practices, functions and the substance of guidance in education. The present chapter utilizes this definition of guidance and urges the development of such systematic conceptual frameworks. The term guidance serv-

\*Dr. Verne Faust, Miami University, and Dr. Lyle D. Schmidt, University of Minnesota, served as consultants to Dr. Gelatt on the preparation of this chapter.

ices will be used without debating its virtues compared to pupil personnel services and other terms.

### *Guidance Research, Theory and Practice*

Disagreement on philosophy and basic objectives for guidance services continued during the last three years. Objectives, when stated, were usually vague, causing confusion and controversy about the roles, function and outcomes of guidance services.

A major research problem for guidance is to demonstrate that something happens as a result of providing such services. Guidance services are often considered to be intangible and, therefore, impossible to appraise. The disagreement about objectives, the confusion over role and the general vagueness of results needs to be overcome by bringing together guidance theory, practice and research. Then, a closer connection can be expected between procedures used and objectives sought. This connection will enable guidance services in schools to become responsive to theoretical principles, scientific investigations and changing needs rather than responding to vague, mystical and intangible notions.

#### *Role and Function*

The American School Counselors Association's official statement on secondary school counselor role and function (Loughary, Stripling and Fitzgerald, 1965) and the official elementary counselor's publication (ACES-ASCA, 1966) did not settle controversy or eliminate disagreement. The publication of these statements probably changed the practice of counselors very little. Although the documents were never intended to achieve such a change, it should be realized that definitions leading to change are needed. Before counselors will change their behavior and before research can assess its effectiveness, counseling must be more clearly defined, specific objectives must be stated, and guidance practices related to those objectives must be identified. A problem with such role statements is rapid obsolescence. Loughary et al. (1965, p. 43) reminded us that "in a relatively short period, several areas of change now apparent could make our currently stated position inappropriate and insensitive to the needs of youth."

Clarke, Gelatt and Levine (1965) presented a decision-making framework for guidance and illustrated how this model could be utilized to evaluate and improve school guidance functions. The authors proposed a strategy for guidance research and suggested three types of research activities: 1) information research which consists of studies collecting data relevant to educational-vocational decisions, 2) evaluation research which identifies program deficiencies, and 3) research which assesses attempts at improvement.

Katz (1966) built a model for guidance role and function which focuses on decision making and which permits the application of modern technology. His model combines three systems of data: a value system, an information system and a prediction system.

*Additional References:* Strowig and Sheets (1967); Grosz (1968).

### *Comprehensive Programs*

Guidance programs are being developed which attempt to present a modern, comprehensive approach to the provision of guidance services. These programs are directly related to the total educational process. Flanagan (1967) is supervising an individual educational development project at the American Institutes of Research. One of the activities of this project is to assemble, develop, test and evaluate guidance procedures for defining a specific educational program for each student. The program will take into account the student's aptitudes, abilities, interests and background. The project is called PLAN: A Program for Learning in Accordance with Needs.

Tiedeman and his associates (Tiedeman et al., 1967; see Chapter 9, this issue) are developing a computer-controlled interactive information system at Harvard for vocational decision-making. Among others, the College Entrance Examination Board and the California Association of Secondary School Administrators are supporting development of a state guidance plan which utilizes a modified systems approach (McDaniel, 1968).

Assessment studies examining elementary and secondary school programs are being developed on a national scale. They are longitudinal research studies generally relating student academic performance and growth with school characteristics and student educational experiences. These studies may certainly influence the future of educational and guidance practices, since what is assessed and made public often becomes the objective given the highest priority in the future. The following are examples of assessment studies: National Assessment Program (Tyler, 1965; Higgins and Merwin, 1967); Equality of Educational Opportunity (Coleman, 1966); Study of Academic Prediction and Growth (Hilton and Myers, 1966); School to College: Opportunities for Postsecondary Education, SCOPE (Tillery, Sherman and Donovan, 1968); and Project TALENT (Flanagan and Cooley, 1966).

An inter professional Research Commission on Pupil Personnel Services (IRCOPPS) has been organized to conduct survey and experimental guidance research throughout the country. Five research centers have been established for this purpose (Goldman, 1966).

*Additional Reference:* Trent and Medsker (1967).

### *Counselor as Researcher*

A counselor's philosophy and practice in guidance will be influenced



by his position on the "clinical vs. statistical prediction" controversy. The degree to which a guidance practitioner is interested in conducting his own research and in the research of others may depend upon his position on the clinical-statistical, subjective-behavioral, intuitive-objective continua.

Goldberg (1968) reviewed the research on clinical judgments published since Meehl's (1954) *Clinical vs. Statistical Prediction*. Goldberg summarized the studies by observing that the typical findings show that clinical judgments tend to be rather unreliable, only minimally related to the confidence and experience of the judge, relatively unaffected by the amount of information available to the judge, and rather low in validity on an absolute basis. Goldberg proposed a learning model (essentially a *research model*) to help clinicians improve their clinical inference ability.

Edgar (1966) examined the underlying philosophical assumptions of guidance research. He called evaluation without values a "wistful wish." If counseling is an orderly and lawful process, then what happens in counseling must lead to verifiable and predictable outcomes. "If counselors are to be purposeful and effective agents of change, they simply cannot remain directionless" (Edgar, 1966, p. 1029).

Chenault (1965) asked for debate on the value assumptions underlying research; she maintained that researchers are handicapped by "the monolithic tradition," that research literature is unilateral, and that professional excellence is narrowly defined. Chenault suggested that "untraditional research" is called for to test but not necessarily to discard traditional assumptions.

Noble intentions and "intangible" effects alone are no longer adequate justification for the existence of guidance practitioners. Practitioners must ask themselves if guidance practices produce desired results.

The need for guidance research to be conducted by local guidance practitioners has recently been acknowledged. The Association of Counselor Educators and Supervisors and the College Entrance Examination Board jointly sponsored a book (Moore et al., 1966), *Research Guidelines for High School Counselors*, to assist counselors in conducting research. The book includes chapters on prediction studies, follow-up studies, opinion surveys, career choice, student attitudes, counseling effectiveness and systems analysis. The American School Counselors Association's Research Committee is currently preparing a practical, descriptive monograph to assist school counselors in conducting local research studies to answer relevant questions about their own schools (Varenhorst, in press).

Eckerson (1967, p. 353), in discussing the overlap in elementary school guidance among eight pupil personnel services, said, "the functions and basic understandings of counselors rest largely on the research and experience of other disciplines." Elementary guidance must, according to Eckerson, equip itself with research methods if it is to develop professional status and carve out its own identity.

Banta (1967) suggested that the new guidance researcher might be a

psychologist concerned with research in contemporary and historical social settings—the community. This future guidance worker will include research as part of his ongoing role, but this research will be in the school and community and not in the sequestered laboratory. This new researcher will have or be affected by several major features. First, there will be a tendency for the research variable to become a value, an issue of vital concern rather than just another convenient variable. Second, the research cannot be an abstract evasion; findings must be applied “judiciously and forcefully” to guidance practices. Third, the new psychologist must have multiple reference groups and give careful consideration to local social conditions and the local power structure in the school and community. Finally, the new researcher, to be successful in this “laboratory in community research,” must develop new “security operations.” He must give up much that he has come to value: the privacy of his office, the rigor of the laboratory, the publication criterion, and the protection of a defined role.

The following points have been generally emphasized in the research literature: 1) a framework for the provision and evaluation of guidance services is desirable; 2) the goals and specific objectives of guidance services need to be explicitly stated; 3) the question of values cannot be ignored; and 4) research that is undertaken must be locally meaningful and related to procedures used and objectives sought.

*Additional References:* Krumboltz (1965); Krumboltz (1966).

### *Guidance Programs*

Concern has increased for K-12 guidance programs. Little is to be gained, therefore, by continuing to emphasize the separation between elementary and secondary school guidance. Such division only fosters the perpetuation of two conflicting camps in school guidance.

#### *Counselor Role and Role Perceptions*

Since most guidance programs are staffed by counselors, what a counselor does and how his role is perceived are important considerations. It becomes increasingly evident, however, that counseling is no longer the only function of a guidance program.

Gabbert, Ivey and Miller (1967) tested the hypothesis that counselors will vary in their ability to work with different types of clients. Client attitudes toward the counselor were examined and found to differ among counselors. Sex differences, type of counseling problem and duration of counseling were found to be important factors relative to client attitudes about counselors.

Perceptions and expectations of the counselor role have been the object of a number of studies. Rosen (1967) reviewed the literature on

preferences of clients with respect to the characteristics and procedures of counselors. He made suggestions concerning methods and areas of needed research, such as more experimental research rather than more correlational survey studies. Further research is needed on the client's preference for counselor personality and counseling procedures, the client's ability to discriminate between preferences and expectations, and the reasons for students' seemingly adverse feelings about discussing personal problems, as compared to educational problems, with high-school counselors.

Pratte and Cole (1965) studied perceptions of referred and self-referred students; they found that voluntary (self-referred) clients had a better understanding of the counselor's role. If it is true that students with a "better understanding" of the counselor's role are more apt to seek counseling, then programs need to explore ways of providing experiences that will enhance such understanding. Of course, the question remains unanswered whether increased seeking of counseling leads to increased counseling effectiveness.

### *Guidance Information and Prediction*

Providing students and adults with information for increased self-understanding, for making predictions, and for expanding alternatives in decision-making is another important aspect of guidance programs. Research can improve the quantity and quality of the information and can increase the effectiveness of presentation methods.

Watley (1966a) studied counselor variability in making accurate predictions by comparing their ability to predict freshman grades. He found that counselors differ significantly and that the ability to predict accurately was not related to the counselor's confidence (Watley, 1966b). Seeking to answer the question: "Do counselors know when to use their heads instead of the formula?", he studied the effect of validation experience on judges known to predict educational criteria with high, moderate and low accuracy (Watley, 1968). Results showed that accuracy did not improve after validation experience; the answer to his question was "No."

Clarke and Gelatt (1967) provided an example of local guidance research designed to improve student decision-making. Results of follow-up research on a graduating class were utilized to assist other students in predicting units needed to enter various types of college. It was found that the type of college entered could be predicted almost as accurately at grade nine as at grade twelve. Clarke and Gelatt also found that many of the poorer students took far more college preparatory courses than they needed for college entrance purposes. Students were able to utilize these data in making educational decisions.

Berdie and Hood (1966) studied predicting plans for college attendance. They concluded that guidance must take into account that the relationships between the predictor variables and the criterion for the

group on the average may be different from these relationships for *each member* of the group. The authors urged, therefore, that guidance programs help each student understand his own background, his own characteristics and his own motives. Students who receive assistance from guidance programs in incorporating group predictive data into their own self-understanding, their environment and their opportunities will increase their ability to make choices.

A person can both remember his past and imagine his future, and since his goals and objectives are affected by his expectations, the self-concept might be described as a series of self-predictions. The importance of providing adequate and relevant guidance information and counseling services on which to base these self-predictions seems evident. In the absence of reliable knowledge, counselors and students will continue to make unwise predictions and decisions.

A systems approach as a framework for the provision of school guidance services is beginning to gain attention and support. (See chapter 9 in this issue.) Cooley (1968), for example, applied the systems approach to guidance in an individualized educational organization. When assisted by the computer, such an approach provides the student with "feedback information" as well as "feed-forward information" to aid him in decision making. Perhaps the systems perspective holds the greatest promise for the future of guidance.

*Additional References:* Tolbert (1966); Bartlett and Green (1966).

### *Research Studies*

Conclusive and relevant research results about school guidance programs were not plentiful in the last three years. The examples presented in this section illustrate the types of research being conducted and the need for a broader approach. Van Hoose and Vafakas (1968) mailed questionnaires to state guidance directors in fifty states, four American territories and the District of Columbia. They received a 100% return. They included questions about the number of counselors employed, their training, their certification status, their source of financial support, and state standards. Van Hoose and Vafakas (1968, p. 538) found that "standards are vague and lack specificity in terms of objectives and producing desired changes in children." Thus, the Van Hoose and Vafakas finding supports the assertion made earlier in this chapter that the variability of titles and the confused perceptions of roles for psychologists, social workers and counselors confounds results and their interpretation.

Barclay (1967) conducted experimental research to determine whether evidence of behavior change could be obtained from a group of fifth-grade children who were subjects of three different treatments. The criteria for change (increased social acceptance) were sociometric scores, teacher ratings, parts of the Holland Vocational Preference Inventory, and a



semantic differential measure. The treatments were: 1) an experimental intervention procedure in one class by psychologist interns, 2) instruction to the teacher on ways to positively reinforce selectively, and 3) a change of teachers after a five-week period. The first treatment included such interventions as sociodrama, a social spelling game and small-group counseling. Results generally favored the first treatment. Barclay urged that more experimental studies instead of correlational surveys be conducted. He called for elementary counselors to function as an integral part of the school rather than functioning as "visiting dignitaries" who appear and summon children to meet with them.

Pierce-Jones, Iscoe and Cunningham (1968) conducted a demonstration-research program in Texas elementary schools. The research study, supported by IRCOPPS, tested the hypothesis that something referred to as "sustained consultation" *about children with teachers* would produce changes in the "mental health promotion orientations" of teachers. This project developed from an interest in seeking alternatives to providing direct guidance services to children. A basic criterion instrument, *Dimensions of Teacher's Opinions* (DOTO), was developed along with other criteria and predictor measures. Results after one year of the experimental treatment revealed that no measurable differences in the orientation of teachers were found on the criterion instruments between teachers receiving consultation and teachers receiving no consultation. Although teachers believed that the child behavior consultation service was valuable and requested its continuation, no evidence of its value was established.

The Pierce-Jones, Iscoe and Cunningham study illustrates two problems often related to the usefulness of research results. The first problem is determining the appropriateness of the criterion. Does it mean anything, for example, to have teachers score higher on the DOTO? Do teachers who score higher *behave* better toward children or teach more effectively? The second problem is fulfilling the need for a variety of criteria. A study such as this should use a variety of criteria, including behavioral evidence. Possibilities for multiple criteria include measures of actual behavior changes in teachers and students and measures of ratings by others of attitude changes in teachers and students. Since teacher consultation was being tested as an alternative to direct service to children, these two treatments should be experimentally compared. Until more research is done with other criteria, it would be unfortunate to conclude that teacher consultation is ineffective. This study provided too little data about what actually happened.

Another IRCOPPS project (Shaw and Rector, 1968) provided a framework for conducting research on school guidance services. This study demonstrated the value of multiple criteria, including behavioral measures. The project emphasized a guidance program as a modification of the school environment through intervention with significant adults. The

participating adults received parent and teacher group counseling. Significant differences were found in favor of students of adult participants vs. non-participants in such areas as school grades; achievement and test scores; and frequency of excused and unexcused absences, tardiness and discipline referrals. The authors suggested the following implications of their study: 1) the model of the prevention of learning difficulties through environmental intervention is feasible; 2) the need for specific objectives and evaluation of outcomes is basic; and 3) the observable behavior of participants and those affected by them must be assessed as well as self-reported attitude change when evaluating outcomes of guidance practices. For example, Shaw and Rector discovered in their research on group counseling that behavior change of participants occurred without apparent attitude change as assessed by a written questionnaire.

*Additional References:* Carkhuff (1966); Sullivan (1966); Kelz (1966).

### *A Personal Commentary*

Based on this reviewer's own experience in school guidance and a survey of three years of literature, these tentative propositions are offered:

#### 1. Non-professional support personnel are necessary.

The debate regarding title, role, function, status and training of guidance workers has been and will continue to be rather futile. Confusion will remain until guidance can clearly identify its purposes and objectives, develop a unifying theory that will suggest procedures, and evaluate both the effectiveness of the procedures and the accomplishment of the objectives.

Present debates on role function and status detract from the capacity of guidance and counseling to meet the current and changing needs of youth. This detraction is caused by what Gordon (1965) called an inappropriate model. In this type of model, a "fully qualified professional person" takes *all* responsibility for guidance and counseling. To combat the inadequacies of these "complete clinicians" to meet current needs, the typically proposed solution is to increase the dosage. That solution solves only the problem of protecting the traditional role identities of present professionals.

Guidance personnel need to recognize that many roles in school guidance can be played by many people with different levels of skills and training. The American Personnel and Guidance Association (1966) has adopted a statement of policy regarding support personnel for the counselor. It is no longer necessary or wise to defend traditional roles and identities. Guidance personnel must attend to their purposes and responsibilities; one of those responsibilities is to conduct comparative experimental studies to answer the question: Who can provide what for whom most effectively and efficiently?

2. Guidance goals and research must be related to the total educational system.

School guidance as a basic construct embracing counseling, school psychology, social work and education should have the general goal of arranging environmental conditions to provide successful learning experiences for all students. This goal would include preventing learning difficulties through environmental intervention and assisting students to become capable and independent decision-makers. Appropriate guidance procedures can be derived and assessed from knowledge of learning processes and from modern theories of social behavior.

One important guidance objective is to identify and respond to current major issues which have psychological, social and educational implications for children. For example, the current personal and social problems facing youth in regard to drug usage, student unrest and protest, and race relations require a vigorous guidance approach. The present problems of education in regard to the relevancy of the curriculum, the appropriateness of evaluation methods, and the need for "affective" learning also demand critical inquiry by guidance researchers. Guidance programs should include active efforts to influence the values, the policies and the practices of schools, based on the knowledge and understanding of student behaviors. Such a change will require guidance workers to become specialists in learning processes, leaders in introducing innovations, and activists in providing in-service training for personnel. It will require the courage to change and to actively ask for change. Guidance research can then be applied to determining to what extent objectives are being achieved, to find out which procedures are most effective in effecting change, and to provide relevant "guidance content" for students and teachers.

3. Guidance information research can contribute to guidance content.

Although the information-giving aspect of guidance enjoys little prestige, it is, nevertheless, a vital activity. Knowledge about decision-making, human judgment, prediction and self-content indicates the importance of information giving. Locally conducted information research can help the content of guidance information become more effective, reliable and relevant. Data from local follow-up studies can help students see alternatives and probabilities for their post-high school decisions. Data from studies on present student interests, abilities and attitudes can help both students and teachers make current decisions. This is not evaluation research, but it is the kind of research that provides data to be interpreted to students and teachers for their use.

4. Guidance research must directly involve students and guidance workers.

Research must be designed and conducted in the schools where the research questions are being asked and where the struggle for relevancy exists. Guidance services and research need to involve students more in determining the kind of services to be offered and to involve both students



and guidance personnel more in conducting research. Involving students in setting their own guidance objectives will assure individually relevant guidance programs and will assist in teacher decision-making. Making students and guidance personnel responsible for participating in studies will improve the chances that research will answer relevant questions, thereby increasing the likelihood that research results will be used.

#### 5. Guidance services and research must be innovative.

Perhaps what is needed most today are new, daring and imaginative approaches to the provision of guidance services and to their evaluation. Cramer and Stevic (1967), reviewing the literature on pre-college guidance research, ably expressed this reviewer's position by stating that current guidance research seems to be "whipping dead horses," neglecting areas that need investigation and using only traditional research models.

New, daring and imaginative approaches may require that some traditional professional identities be ignored and that some mystical and intangible goals be translated into more manageable and concrete objectives. For example, a systems approach will require specificity of objectives. A variety of personnel and a variety of outcome criteria are needed in research. More careful listening to students and teachers is required so guidance services and research can be conducted with those students and teachers rather than merely providing services for them.

Guidance personnel must provide leadership for the continued development of new school organization and curricula; they can no longer be mere spectators. The flexibly scheduled school, the non-graded school, the extensive use of independent study, individualized instruction, and the greater freedom and responsibility of the student in making decisions about his education all have profound implications for guidance research and practice.

Participating in the creation of new educational institutions necessitates new kinds of guidance research. Such research will provide not only a measurement of change in our institution, but will furnish an avenue for the development and communication of change.

### Bibliography

- ACES-ASCA Joint Committee on the Elementary School Counselor. Report of the Committee on the Elementary School Counselor. Washington, D. C.: ACES-ASCA, 1966. (Mimeo.)
- APGA, Support Personnel for the Counselor: Their Technical and Non-Technical Roles and Preparation. *Personnel and Guidance Journal* 45: 858-861; 1967.
- Banta, Thomas J. "Two Voices of Psychology and the New Psychologist." *Journal of Educational Research* 61: 22-23; September 1967.
- Barclay, James. "Effecting Behavior Changes in the Elementary Classroom." *Journal of Counseling Psychology* 14: 240-247; 1967.
- Berdie, Ralph F. and Hood, Albert B. "How Effectively Do We Predict Plans for College Attendance?" *Personnel and Guidance Journal* 44: 487-493; 1966.
- Chenault, Joann. "Research and the Monolithic Tradition." *Personnel and Guidance Journal* 44: 6-10; 1965.



- Clarke, Robert and Gelatt, Harry B. "Predicting Units Needed for College Entrance." *Personnel and Guidance Journal* 46: 275-282; 1967.
- Clarke, Robert; Gelatt, Harry B.; and Levine, Louis. "A Decision Making Paradigm for Local Guidance Research." *Personnel and Guidance Journal* 46: 275-282; 1967.
- Coleman, James S. *Equality of Educational Opportunity*. U. S. Office of Education, Dept. of Health, Education and Welfare. Washington, D. C.: Superintendent of Documents, Gov't. Printing Office, 1966. (ED 0 12 275.)
- Cooley, William W. *Computer Systems for Guidance*. Invited Address to Division E, Amer. Educ. Res. Assoc., February 9, 1968. Pittsburgh: School of Education, Univ. of Pittsburgh. (Mimeo.)
- Cramer, Stanley and Stevic, Richard R. "What's New in Pre-college Research?" *College Board Review* 65: 24-29; Fall 1967.
- Eckerson, Louise Omwake. "Realities Confronting Elementary School Guidance." *Personnel and Guidance Journal* 46: 350-354; 1967.
- Edgar, Thomas E. "Wistful Wish: Evaluation Without Values." *Personnel and Guidance Journal* 44: 1025-1029; 1966.
- Flanagan, John C. and Cooley, William W. (editors). *Project TALENT One Year Follow Up Studies*. Cooperative Research Project No. 2333. Pittsburgh: Univ. of Pittsburgh, 1966.
- Flanagan, John C. "Functional Education for the Seventies." *Phi Delta Kappan* 49: 27-31; 1967.
- Gabbert, Karen H.; Ivey, Allen E.; and Miller, C. Dean. "Counselor Assignment and Client Attitude." *Journal of Counseling Psychology* 14: 131-136; 1967.
- Goldberg, Lewis R. "Simple Models or Simple Processes? Some Research on Clinical Judgments." *American Psychologist* 23: 483-496; 1968.
- Goldman, Leo. "IRCOPPS: A Report of Progress." *Personnel and Guidance Journal* 45: 402-406; 1966.
- Gordon, Jesse E. "Project Cause, Some Implications of Subprofessionals' Training." *American Psychologist* 20: 334-343; 1965.
- Higgins, M. J. and Merwin, J. C. "Assessing the Progress of Education." *Phi Delta Kappan* 48: 378-380; 1967.
- Hilton, T. L. and Myers, A. E. *Growth Study II*. Research Bulletin 66-5. Princeton, N. J.: Educational Testing Service, February 1966.
- Katz, Martin. "A Model of Guidance for Career Decision-Making." *Vocational Guidance Quarterly* 15: 2-10, 1966.
- Kehas, Chris D. "Theoretical Formulations and Related Research." *Review of Educational Research* 36: 207-218; 1966.
- Loughary, John W.; Stripling, Robert O.; and Fitzgerald, Paul W. (editors). *Counseling, A Growing Profession*. Washington, D. C.: Amer. Personnel and Guidance Assoc., 1965.
- McDaniel, Henry B. *The California Guidance Project*. Palo Alto: College Entrance Examination Board, Western Regional Office, October 1968. (Mimeo.)
- Meehl, Paul E. *Clinical Versus Statistical Prediction*. Minneapolis: Univ. of Minn. Press, 1954.
- Moore, Gilbert D. et al. *Research Guidelines for High School Counselors*. New York: College Entrance Examination Board, 1967.
- Pierce-Jones, John; Iscoe, Ira; and Cunningham, Grover. *Child Behavior Consultation in Elementary Schools: A Demonstration and Research Program*. Austin: IRCOPPS Personnel Services Research Center, Univ. of Tex., 1968.
- Pratte, Harold E. and Cole, Charles. "Source of Referral and Perception of the Counselor." *Personnel and Guidance Journal* 44: 292-294; 1965.
- Rosen, Albert. "Client Preferences: An Overview of the Literature." *Personnel and Guidance Journal* 45: 785-789; 1967.
- Shaw, Merville C. and Rector, William H. *Modification of the School Environment Through Intervention with Significant Adults*. Western Regional Center of the Interprofessional Research Commission on Pupil Personnel Service, Monograph No. 7. Chico, California: Chico State College, 1968.
- Tiedeman, David V. et al. *An Information System for Vocational Decisions*. Annual Report, 1967. Cambridge, Mass.: Harvard Univ., 1967.
- Tillery, Dale; Sherman, Barbara; and Donovan, Denis. *Helpfulness of Parents, School Personnel and Peers to Students with Different Educational Aspirations*. Paper

- presented to APA Convention, San Francisco, 1968. Berkeley: Center for Research and Development in Higher Education, Univ. of Calif.
- Tyler, Ralph W. "Assessing the Progress of Education." *Phi Delta Kappan* 47: 13-16; 1965.
- Van Hoose, William H. and Vafakas, Catherine. "Status of Guidance and Counseling in the Elementary School." *Personnel and Guidance Journal* 46: 536-539; 1968.
- Varenhorst, Barbara (editor). *The Spirit of Inquiry: Using Research as a Practical Aid in Guidance*. Washington, D.C.: Amer. Personnel and Guidance Assoc., in press.
- Watley, Donovan J. "Counselor Variability in Making Accurate Predictions." *Journal of Counseling Psychology* 13: 53-61; 1966a.
- Watley, Donovan J. "Counselor Confidence in Accuracy of Predictions." *Journal of Counseling Psychology* 13: 62-67; 1966b.
- Watley, Donovan J. "Do Counselors Know When to Use Their Heads Instead of the Formula?" *Journal of Counseling Psychology* 15: 84-88; 1968.

### Additional References

- Bartlett, Claude J. and Green, Calvin G. "Clinical Prediction: Does One Sometimes Know Too Much?" *Journal of Counseling Psychology* 13: 267-270; 1966.
- Carkhuff, Robert. "Counseling Research, Theory and Practice—1965." *Journal of Counseling Psychology* 13: 467-480; 1966.
- Grosz, Richard D. "Effect of Client Expectations on the Counseling Relationship." *Personnel and Guidance Journal* 46: 797-800; 1968.
- Kelz, James W. "The Development and Evaluation of a Measure of Counselor Effectiveness." *Personnel and Guidance Journal* 44: 511-516; 1966.
- Krumboltz, John. "Behavior Counseling: Rationale and Research." *Personnel and Guidance Journal* 44: 383-387; 1965.
- Krumboltz, John. *Stating the Goals of Counseling*. Monograph No. 1. California Counseling and Guidance Association, 1966.
- Strowig, R. Wray and Sheets, Stanley E. "Student Perception of Counselor Role." *Personnel and Guidance Journal* 45: 926-931; 1967.
- Sullivan, Howard J. "The Effects of Counseling as a College-Encouragement Procedure." *Journal of Counseling Psychology* 13: 93-97; 1966.
- Tolbert, E. L. "Predictability in Testing and Counseling." *Personnel and Guidance Journal* 45: 219-226; 1966.
- Trent, James and Medsker, Leland. *Beyond High School*. Berkeley: Center for Research and Development in Higher Education, Univ. of Calif., 1967.

### AUTHOR

Gelatt, Harry B. Address: Palo Alto Unified School District, Palo Alto, California  
 Title: Director of Guidance Age: 42 Degrees: A.B., San Jose State; M.A., Stanford Univ.; Ed.D., Stanford Univ. Specialization: School Guidance Programs, Elementary and Secondary.



### 3: HIGHER EDUCATION PROGRAMS AND STUDENT DEVELOPMENT

Harold A. Korn\*

*Stanford University*

That higher education is facing a time of crisis and change is generally conceded, but agreement ends with that concession. Finding solutions to the problems is complicated by competing, often antagonistic views. In this chapter, focused on recent literature and research on the student in higher education, an attempt is made to clarify the problems, identify relevant conceptual and methodological approaches, and raise questions concerning the most effective path to constructive change. The period from June 1965 to October 1968 is reviewed, with occasional reference made to earlier studies.

"Higher education programs" usually refer to those activities which take place outside the formal curriculum. These activities are performed or supervised by members of the student personnel services professions. Traditional student personnel programs need critical reappraisal in light of the current needs of students. "Student development" refers to a point of view which is consistent with the goals of the traditional liberal arts philosophy, but goes an important step further by recognizing that the affective domain of behavior must be explicitly taken into account along with the cognitive domain. Unless both of these areas of behavior are

---

\*Dr. Alexander Astin, American Council on Education, Dr. Clyde Parker, University of Minnesota, and Dr. Nevitt Sanford, Stanford University served as consultants to Dr. Korn on the preparation of this chapter.



recognized in teaching the curriculum and in the activities of the extra-curriculum, the goals of a liberal education cannot be achieved.

In this chapter, the author attempts to demonstrate the relationship between the findings of the behavioral sciences and the practices of student personnel services. The author also strives to center the attention of behavioral scientists upon the pressing day-to-day problems of the campus community.

## *Higher Education: A Social System Interrelated with Other Social Systems*

### *Higher Education and Contemporary Society*

One hundred years ago students were the most effective agents of campus reform. Rudolph (1965), in making this point, suggested that earlier students were able to use colleges as an instrument of their maturation; now the university serves professors more effectively than students.

Few issues have remained the same during the hundred-year span that Rudolph sought to understand; one of those unchanged issues is the yearning of students for a sense of purpose and fulfillment. For decades, higher education programs have been philosophically committed to helping students develop. But as Koile (1966) phrased it, the profession has been "Forever the Bridesmaid." The student personnel profession's relatively minor influence on the functioning of higher education must be understood in the context of higher education's role in society. A perspective on what has happened in our society is necessary, even though some may question the relevance of history and social criticism to scientific research.

In the past 100 years, higher education has moved from a numerically-small and only marginally-powerful social institution to the center stage of what economist and social critic Galbraith (1967b, see especially chapter 25) calls the New Industrial State. The Academic Revolution, born of the increasing demand for specialized talent for our complex technology, was examined in detail by Jencks and Riesman (1968). They argued that the prestige and power of graduate schools with their discipline-centered programs have led to the stifling of liberal arts goals. Specialization and bigness are not inherently antithetical to human values and personal development, but they tend to be inversely correlated.

The threat of technology to the human condition has attracted the attention of numerous social critics: Marcuse (1964) and Mumford (1966) are among the most articulate. In *Beyond Alienation*, Becker (1967) has presented a provocative analysis of the current educational dilemma and an imaginative set of alternatives for curriculum reform based on the "New Moral View of the World."

It is necessary to move from the abstract discussion of the conditions of contemporary man to a detailed analysis of how our social institutions influence the lives of individuals. In the past, the emphasis has too often been either on understanding society or understanding the individual, rather than on their mutual interaction. A few men of wisdom, however, have offered a perspective on this interaction; an outstanding example is found in Sanford's (1964) analysis of higher education.

### *Systems Perspective*

In the last decade, the systems perspective has been gathering enthusiasts because it offers, in schematic form, both conceptual frameworks and methods of inquiry adequate to the task of understanding exceedingly complex phenomena. Boguslaw (1965) defined a systems perspective as a way of recognizing that performance in any one of a series of successive stages is dependent upon performance in other stages. This interdependence requires that every stage in a sequence be carefully analyzed and its relationship to other stages be painstakingly spelled out. Churchman (1968) argued that the technology of systems science can be used to humanize education by forging a precise and detailed description of the existing educational system so actual operations and output can be evaluated relative to goals.

In *Integrating the Individual and the Organization*, Argyris (1964) explored ways of achieving organizational effectiveness and individual development. His discussion of the unintentional consequences of certain organizational procedures and the individual's quest for *sub rosa* means of not conforming to the organization or of covertly changing it is pertinent to higher education.

Katz and Kahn (1966) presented a new approach to the study of behavior within organizations in *The Social Psychology of Organizations*. They utilized open systems theory to stress the necessary dependence of any organization upon the individuals who are members of the organization—and most important, upon the willingness of the individuals to perform as members. The need to view higher education within a systems perspective is illustrated in the next chapter section, entitled *The System in Trouble*. In that section, the author suggests that the current crisis in higher education has resulted from the failure to consider the consequences of various systems components, such as administrative procedures, educational policies, and environmental pressures. Evidence that the system is out of balance should encourage a search for system-wide answers rather than isolated emergency remedial actions.

*Additional References:* Barton (1961); Boocock (1966); Buckley (1968); Davis (1967); Pervin (1967); Thoresen (in press).

## *The System in Trouble*

### *Academic Performance and Real-Life Success*

The Grade Point Average (GPA) is to the educational world what Gross National Product (GNP) is to the economic world. GPA and GNP are convenient indices which have assumed properties bordering on the mystical. A careful analysis of the consequences of using both indices so pervasively to evaluate success could lead to significant changes in the way social institutions are organized.

Berle (1968) introduced a companion concept to Gross National Product called the Gross National Disproduct. He examined whether the creation of GNP might involve as much frustration as satisfaction. This point is important as a commentary on the value dilemmas of society and as an indication of how a nationally accepted index of success may hide more than it reveals. GPA has been as uncritically accepted in the educational world as GNP was accepted in the economic world. Some institutions have responded to student pressure by changing from letter or numerical grading to a system of descriptive adjectives; others have experimented with pass/fail criteria. There has not been adequate recognition, however, of the need to examine the purposes, interrelationships, and hidden consequences of grades or any set of performance criteria.

Lavin (1965), in *The Prediction of Academic Performance*, provided a thorough introduction to this complex issue. He analyzed studies to determine the validity of ability, personality and sociological variables as predictors of GPA. None accounted adequately for the variance in GPA; seldom was more than 50 per cent of the variance accounted for. A multivariate approach is necessary to predict GPA, multivariate in the broad sense of including various measures of the social environment. Such an approach should encompass student-teacher relationships and the influence of informal peer-group norms upon academic achievement.

GPA, however, is a very ambiguous and questionable criterion. Hoyt (1965) questioned how adequately grades predict future success. After conducting an extensive review, Hoyt concluded that college grades have no more than a very modest correlation with adult success, no matter how success is defined. He also provided a valuable analysis of the meaning of college grades and the whole problem of evaluation in higher education. Hoyt suggested the use of a profile of student growth and development instead of grades. In a later study, Hoyt (1968) also considered the problem of forecasting success in specific colleges.

A commitment to the development of the entire individual has always been implicit in the philosophy of a liberal education. Despite this commitment, there is an almost exclusive use of grades to measure success of the individual and, indirectly, the success of the educational system. However, as Holland and Richards (1965) have demonstrated, academic

potential and achievement have very little relationship to other kinds of non-academic and socially important performance. Richards and Lutz (1967) found that non-academic accomplishments, such as leadership, could be assessed with moderate reliability and could be predicted.

*Additional References:* Bowers (1967); Juola (1968); Webb (1967).

### *The College Dropout*

In a nation increasingly committed to the ideal of higher education for all, the lack of congruence between the individual and the system is an inevitable danger. The manpower demands of our technological society, a sharp increase in the number of teenagers, and commitment to advanced study have forced some states to develop master plans for higher education. Knoell (1966) described the efforts made in New York and California to adapt a social institution designed for the elite to the requirements of a democracy.

The lack of congruence between the needs of the individual and the demands of a system of higher education is evidenced by the number of dropouts. At one time or another, 50 per cent of the college population drop out; however, this percentage includes those students who transfer or later return to the same school. The issue of "the college dropout and the utilization of talent" was carefully explored in a book by that title, edited by Pervin, Reik and Dalrymple (1966). These authors indicated that dropping out may be as beneficial to one student as it is detrimental to another. In spite of this, an inescapable social attitude against dropouts persists: individuals who cannot conform to the usual pattern of a four-year college curriculum are treated with disdain.

Ford and Urban (1966, p. 106) broadened the perspective on the utilization of talent by suggesting that "the phenomena of the college dropout implies a basic flaw in our entire educational structure. . . . We need to invent forms of education that are appropriate for those who are not highly effective as symbolic learners."

Heist (1967) examined the problems and needs of creative college students. He found that students identified as creative leave college more often or just as often as those who are not so identified.

*Additional References:* Axelrod (1968); Pallet and Hoyt (1968).

### *The Black Student*

For black students, the discrepancy between individual needs and institutional requirements reaches tragic proportions. For the white student the curriculum appears to have face validity; the average white student has been taught to believe that he should find the history of Western Civilization, English literature, etc. interesting. Many black students find this an unwarranted assumption and are demanding more courses that



emphasize Afro-American culture. Some educators find hope intrinsic in the very problem. "These ghetto youth," Bush (1967, p. 2) stated, "challenge traditional and accepted academic procedures, inspire the middle-class students to venture beyond pat answers and normal study patterns, and cause professors to reevaluate present teaching methods."

In *Study of Negro Students at Integrated Colleges*, Clark and Plotkin (1963) found that GPA was related to fathers' occupations. They also found that black students had a lower dropout rate than whites at the same colleges. Bradley (1967) found that American College Test (ACT) scores were not equally predictive of GPA for black and white students in predominantly white state colleges and universities in Tennessee. Bradley saw the problem in terms of a need for interracial education at elementary and secondary school levels. Stanley and Porter (1967) found, however, that the curriculum of the Scholastic Aptitude Test with college grades was about the same for white students in white colleges and Negro students in Negro colleges. This similarity in predictability occurred despite the very large difference in mean aptitude scores between the white and Negro students.

McClain (1967) studied the personality characteristics of black college students in the South. Using Cattell's 16 Personality Factors inventory, McClain found marked differences between black and white students in dominance and other aspects of personality development. Boney (1967) explored a remedial approach to the problem and proposed a model for assertion training in academic situations. The efforts and frustrations of white institutions in recruiting black students were graphically described by Sabine (1968). He quoted, for example, a conversation which took place on a basketball court, detailing how the recruiter talked to a prospective black student while he was practicing set-ups.

Just as the system of higher education must be understood within the framework of contemporary society, so must the black student be understood in the context of that society which has long ignored him. A book of essays edited by Parsons and Clark (1966) spelled out the dimensions of the Negro-American experience. At the present time the most significant dimension can be described in terms of the polar concepts *integration* and *separatism*. The long-sought integration into the mainstream of American society now seems to many black Americans to be too slow and too filled with disappointments. A sense of identity and a basis for self-respect are now being pursued through a visible Afro-American culture. Higher education, unprepared to deal with the culturally disadvantaged student in its traditional curriculum, is now being bombarded with demands for a new kind of curriculum. In one sense, the response by higher education to the challenge presented by the black students will be the key to its readiness to respond to a wide variety of problems.

*Additional References:* Baratz (1967); Clark (1967); Dreger and Miller (1968); Harris and Reitzel (1967); Harrison (1967).

### *Psychological Needs and Impersonality*

The dreams, accomplishments and failures of mankind are the subject matter of a liberal education. Many students expect that exposure to the wisdom of the past and the accumulated knowledge of the present will promote their personal development. When this fails to happen, these students become apathetic, rebellious or alienated. Pervin (1966) considered the reality and non-reality in student expectations of college; he found that students are unrealistic in expecting the college to emphasize abstract understanding or social welfare.

Katz and associates (1968) presented the results of a five-year longitudinal study of changes in students at Stanford and the University of California at Berkeley. The student included semi-annual interviews for selected samples of 250 students, an analysis of freshman-senior personality scales, and a variety of case studies. The researchers found that college had little impact on student development. This finding led the researchers to conclude that the answer to the present dilemma lies in changing the curricular offerings and other components of the educational environment. The authors proposed that the focus of education should be on the student and his development, rather than on the accumulation of course credits. This proposal would involve keeping a profile of the student's development—not merely keeping a transcript of his grades.

Trent and Medsker (1968) studied ten thousand high-school graduates to determine what impact attending college or going to work has on individuals. The authors concluded: college seems to foster the growth of autonomy and intellectual disposition of its graduates only in comparison with those of their peers who had less college or none at all. Evidence collected from questionnaire and interview items suggested that most college graduates are apathetic to intellectual inquiry and social issues.

Both the Katz (1968) and the Trent and Medsker (1968) studies focused on how higher education is thwarting the psychological needs of many students. Some of these needs seem mundane and reflect the students' desire for work assignments that are relevant instead of being mere hurdles to jump. Other student needs are thwarted by administrative philosophies like *in loco parentis*, which is antithetical to late adolescents' wish for freedom. If attention is paid to these system details which distress the student, perhaps student motivation can be changed from fighting the system to more active cooperation with the system.

Once changes are made in the ways students are thwarted, it is possible to move into the realm of what man can become if he is encouraged to use his full human potential. That higher education should be centrally involved in this endeavor was clearly articulated by Sanford (1967) and

Freedman (1967). Kauffman et al. (1968), in their succinct explication of this position, described the psychological needs of students in higher education and proposed a broad, heuristic theory of human development; they (1968, p. 58) argued that "the quality of human relationships in higher education must be improved not simply because it will enable students to spend happy and more fulfilling years in college or because many of the present conditions in higher education are intolerable, but primarily because, unless the trends toward giantism and dehumanization are reversed, the college will not be able to educate even the technician."

### *Student Unrest and Radical Students*

It is both ironic and inevitable that students are the ones forcing a re-examination of the system of higher education. Ironic because the presumed wisdom and knowledge of student personnel, professionals and related behavioral scientists have not had the same impact. Inevitable because students suffer most from the inequities of the system and are in a position to gain the most while risking little. Potter (1965, p. 71), a student, gave some insight into the students' rhetoric. Potter wrote, "If my tone is not detached, it is because I find with my compatriots one source of concern—a world in which men so easily detach themselves emotionally from the things they create."

Keniston (1968) analyzed the power of this "youth culture" to effect change; he attributed some of its potential force to contradictions inherent in society. Great economic wealth permits the development of a class of students and young adults free to pursue idealistic goals, but the very factors producing this wealth also incur conditions antithetical to human values, thus inciting idealists to seek change. Flacks (1967) also found student protest based on deep discontent among certain high-status youth.

Sampson (1967) explored how this youth culture evolved from the student apathy of the 1950's. That activist students are intelligent and productive seems to be generally agreed (Katz, 1968). However, there is no clear explanation of why students are activists in the sixties and their "brothers" were not in the fifties. The Vietnam War and the Civil Rights movement appear to be among the causes of activism. Although student protestors are by no means in the majority, Gales (1966) found that 80 per cent of the 400 students randomly selected and interviewed at Berkeley favored the goals of the Free Speech Movement. However, the students in that 80 per cent group did not all agree with the FSM methods of attaining their goals.

Activists find all social institutions defective, but they concentrate upon the institutions of higher learning. The tradition of academic freedom makes higher education particularly vulnerable to the accusation of abridging students' constitutional rights. Williamson and Cowan (1966) surveyed a representative sample of American institutions of higher educa-



tion and found that students' freedom of expression was subject to a variety of definitions. Problems posed by these varying definitions were analyzed by Armacost (1966).

A focal point of student discontent is university governance. A penetrating analysis and a far-sighted set of proposals were recently put forth by the Faculty Student Commission on University Governance at the University of California. Two major themes were explored: promoting a greater degree of open discussion in a spirit of mutual trust, and decentralization of decision-making powers to smaller campus units sharing common purposes.

## *Student Growth Through Liberal Education*

### *Studying the Impact of College*

Newcomb and Feldman (1968) made a substantial contribution to the study of college students by reviewing extant literature on the impact of college. The authors posed the following question: Under what conditions have what kinds of students changed in what specific ways? They concluded that an increase in "open-mindedness," a decrease in conservatism concerning public issues, and a growing sensitivity to aesthetic and inner experiences takes place during college. Newcomb and Feldman warned against the pitfalls of cross-sectional studies and advocated use of longitudinal data. They also examined in some detail the tendency to mask individual changes by presenting results only in terms of mean differences. These authors concluded that college experience, at the very least, accentuates changes inherent in the developmental patterns of individuals.

A fundamental problem for studies using objective personality scales to measure change is centered on the validity of such measures. Korn (1968) used a cluster analysis technique to study empirical dimensions in the structure of individual scales (the Social Maturity and Impulse Expression Scales of the Omnibus Personality Inventory). The multidimensional structure of such individual personality scales compounds the difficulty of interpreting what is actually represented by the changes in mean scores.

The conceptual confusion that surrounds the concepts of personality development and socialization is a basic problem in studying change during college. The use of so-called personality scales to measure change resulting from college experience inadvertently implies change in basic dimensions of personality, whereas the change may actually be the result of the socializing influence of prevailing peer group attitudes. Unfortunately, behavioral scientists who use socialization as a primary theoretical construct do not use personality constructs in their theories, and vice versa. Relating these two approaches to understanding human behavior could be a fruitful new area of research on college students. Brim and Wheeler



(1966) made a valuable contribution by delineating how socialization occurs after childhood. Madison (1966) used the study of personality in college students to examine new theoretical constructs in personality theory. Using extensive in-depth interviews, he explored the consequences of one's "child-self" in continuing to influence decisions of later life.

*Additional References:* Barton (1968); Dressel (1965a); Harris (1963); Plant (1965).

### *Critical Thinking, Open-mindedness, and Resistance to Change*

A basic assumption of the liberal arts curriculum is that critical thinking and open-mindedness are qualities of human behavior which can be encouraged. A further assumption is that these qualities of human behavior are eagerly sought by students and faculty. There is an extensive literature in the behavioral sciences that indirectly bears on these assumptions. It is remarkable, however, that these basic assumptions have not been directly tested. It is equally remarkable that behavioral scientists, who have been most directly concerned with critical thinking and open-mindedness, have not seen fit to study the great natural experiment created by the liberal arts curriculum.

A pioneering study was the Newcomb et al. (1967) analysis of the long-term results of a Bennington College education in the 1930's. Using a combination of interviews and questionnaires, the researcher found that social awareness, a major objective of the Bennington curriculum, persisted 25 years later. The persistence of these attitudes was related to the social environments that these women had created for themselves after leaving college. By choosing husbands, friends and co-workers with similar social philosophies, the attitudes which these women developed earlier were reinforced throughout the subsequent years.

Unfortunately, few studies have focused directly on the influence of the college experience on open-mindedness and critical thinking. Rokeach (1968) moved toward a less authoritarian view of the nature of belief systems and how they are influenced. This area of research has important implications for higher education because it focuses on the human tendency to adhere to any established belief system. It may be that students do want to examine critically their own belief systems. It should be acknowledged, however, that this examination may be a painful emotional experience; students will find ways to avoid becoming genuinely open-minded and self-critical.

The student personnel professional who is concerned with finding ways to foster critical thinking and open-mindedness can make a genuine contribution to the liberal education of students. Korn (1966), for example, called for collaboration between professors who could teach content and other professionals who could take into account the various sources of resistance which students bring to the liberal arts curriculum. Thoresen

(1966a, b) discussed several techniques for exploring this kind of collaborative effort. Thoresen used both video tape in the classroom and a study of the characteristics of non-participants in discussion sections to focus attention on the affective domain in higher education.

*Additional Reference:* Dressel and Lehmann (1965).

### *New Approaches to Implementing Traditional Goals*

Innovative thinking and some experimentation have evolved from seriously questioning the system of higher education. A recent Western Interstate Commission for Higher Education institute resulted in a volume of essays (Minter, 1967) which surveyed much of the recent effort to personalize higher education. Cluster colleges and experimental programs are two broad categories of innovation.

#### *Cluster College*

In an effort to combat the impersonality of the multiversity, smaller semi-independent units called cluster colleges are being created. Martin (1968) indicated, however, that even cluster colleges were developing along disciplinary lines. Martin called for programs focused on urban problems and set in strategically located city buildings.

The promise inherent in the idea of cluster colleges is evident; equally clear is that the implementation of the idea presents many problems, as summarized by Kells and Steward (1967). These problems run the gamut from petty departmental jealousies to fundamental questions relating to philosophy of education. Astin (1967), a participant in the Conference on the Cluster College, focused attention on the need to experiment with the composition of student groups participating in a cluster-college program. Astin called for longitudinal evaluative studies that could systematically assess the impact on students of different types of cluster colleges.

The cluster-college setting is ideal for activating what has been learned about the influence of peer groups on student development. Ideas found in Newcomb and Wilson (1966), an excellent introduction to the study of college peer groups, might be implemented in cluster colleges. Newcomb (1967) stressed the importance of an internal organization in the college that would lead to the formation of meaningful student-faculty groups and involve both their personal and academic lives.

#### *Experimental Programs*

In addition to the development of cluster colleges, various experimental programs give expression to the need for change. Nixon (1967), a student who helped develop the Experimental College at San Francisco State College, found that one significant effect of the program was that

students became alert to the possibility of change. The students no longer blindly assumed that education had to be presented as they were receiving it in their classrooms.

A product of student protest at the University of California at Berkeley was the report of the Select Committee of the Academic Senate (University of California, 1966), informally known as the Muscatine report. In this report, the committee asked for a variety of reforms in the structure and organization of the University; the reforms included the establishment of a Board of Educational Development that has the power to create new instructional programs and by-pass the established University machinery.

Axelrod (1967, p. 328), finding that the "problem" is often formulated prematurely in new programs, proposed a curricular plan which "attempts to help the student free his mind of the very middle-class bias which determines, more often than not, how a societal problem is perceived and hence formulated."

McKeachie (1967) called for further research on the characteristics of college teachers, with particular reference to their capacity to see course material from the student perspective. Acker, Danskin and Kennedy (1967) illustrated how the knowledge of student characteristics could determine the structure of the curriculum. They found that the students entering the College of Agriculture expected and needed course work that would relate to their past experiences and future goals.

At a time when students are demanding relevance, "games" may appear to be out of place. However, Boocock and Schild (1968) have offered convincing arguments for using simulation games in learning. Issues and processes from life can be abstracted, permitting students to examine the consequences of different strategies in decision-making.

*Additional References:* Axelrod (1967); Dressel (1965b).

### *Directions for the Future*

A systems perspective requires that the consequences of all actions be examined in light of avowed goals. When the student personnel profession considers both the ineffectiveness of past actions and the unsolved current problems, a dramatic reappraisal of the professional role seems inevitable. The urgency of problems facing higher education and the paucity of relevant knowledge available from behavioral sciences demonstrate that a program of action research is vital. Sources of student unrest must be dealt with by developing programs which come to grips with impersonality and the perceived irrelevance of large parts of the curriculum. Such programs must be established in ways which permit critical evaluation of program success and failure. Fairweather (1967) has made an interesting beginning toward setting up a community-wide program of social inno-

vation; he has also described a system for evaluating the success of his program. Webb et al. (1966) described several ways to obtain unobtrusive measures of social behavior that may be useful in evaluating change.

The need to respond to pressing immediate problems is surpassed by the even greater need to develop a socio-psychological theory adequate enough to increase the understanding of the structures and processes of higher education. Researchers must recognize the importance of developing a science of human behavior that will encompass the complexity and scope of the long-range problems of higher education. A number of studies (typical in many ways of those generally found in professional journals of higher education) were excluded from this review because they lacked scope. Too many studies were limited to correlating the assessment variables of a certain instrument with some other criterion. The researcher neglected to criticize either the validity of the assessment measure, the meaningfulness of the criterion, or the significance of the question being pursued. This lack of critical appraisal made replication studies by other researchers difficult—often the work was not even followed up by the original researcher. Unfortunately, experimental studies in which significant variables were systematically manipulated were exceedingly rare.

Although it may be relatively easy to say that certain kinds of studies should be discouraged because their only contribution is to the bibliography of the author, it is more difficult to say what studies should be done. At the present stage of development in the behavioral sciences, no one can claim a corner on truth. However, it is possible to point to major research projects which, though very different in character, are models of what research efforts should be. One excellent example is the path charted by Astin and his colleagues (1966, 1967, 1968); they have developed a research program and an accompanying methodology to describe graphically what is actually happening in many sectors of higher education. Their evidence suggests that traditional indices of institutional quality do not measure variables which contribute to student achievement. Such empirical evidence must be sought to provide a clear picture of what is actually happening. Problems associated with this empirical research strategy were described by Creager and Astin (1968) and Astin (1964).

Another style of research cannot be so easily characterized, since it deals with what should be as well as what is. A statement of values and a theory for translating them into action programs are implicit in any discussion of what ought to be. In his recent work, Bloom (1968) challenged many fundamental assumptions about the educational process by advocating "learning for mastery." Bloom suggested that most students can master most material; he expressed the need for the entire educational system to become student oriented—to accept the burden of finding ways and means of becoming effectual rather than covertly blaming students unable to learn from the traditional prescribed procedures.



The research efforts by Astin and Bloom can be viewed as two kinds of input to a systems approach to higher education. As the picture of what is actually happening in higher education becomes clearer, and as agreement grows concerning what should be happening in higher education, counseling will be faced with the enormous challenge of implementation. The role of the student personnel profession will become clearly defined, because it alone has the potential for acquiring the needed comprehensive knowledge of individual student characteristics and the commitment to create a campus environment conducive to mastery learning. By combining the student development point of view with the educational philosophy implicit in the concept of mastery learning, higher education can find a path out of its current crisis.

### Bibliography

- Acker, D. C.; Danskin, D. G.; and Kennedy, C. E., Jr. "Student Characteristics in Curriculum Planning." *Journal of College Student Personnel* 8: 381-384; 1967.
- Argyris, C. *Integrating the Individual and the Organization*. New York: John Wiley, 1964.
- Armstrong, Peter. "The American Association of University Professors and the Association of American Colleges Statements on Student Freedoms—A Comparison and Discussion of Viewpoints." *Proceedings of the NASPA*. Seattle: The Association, June 1966.
- Astin, Alexander W. "Criterion Centered Research." *Educational and Psychological Measurement* 24: 807-822; 1964.
- Astin, Alexander W. "The Conference on the Cluster College Concept: Students." *Journal of Higher Education* 38: 396-397; 1967.
- Astin, Alexander W. *The College Environment*. Washington, D.C.: American Council on Education, 1968.
- Astin, Alexander W.; Panos, Robert J.; and Creager, John A. *A Program of Longitudinal Research on the Higher Educational System*. American Council on Education Research Reports 1, No. 1. Washington, D. C.: the Council, 1966.
- Astin, Alexander W.; Panos, Robert J.; and Creager, John A. *Implications of a Program of Research on Student Development in Higher Education*. American Council on Education Research Reports 2, No. 2. Washington, D. C.: the Council, 1967.
- Axelrod, J. "An Experimental College Model." *Educational Record* 48: 327-337; 1967.
- Becker, E. *Beyond Alienation*. New York: George Braziller, 1967.
- Berle, A. A., Jr. "What G. N. P. Doesn't Tell Us." *Saturday Review* 51, No. 35: 10-12; 1968.
- Bloom, B. S. *Learning for Mastery*. Report of UCLA Center for the Study of Evaluation of Instructional Programs 1, No. 2. Los Angeles: Univ. of Calif., May 1968.
- Boguslaw, Robert. *The New Utopians. A Study of System Design and Social Change*. Englewood Cliffs, N. J.: Prentice-Hall, 1965.
- Boney, J. Don. "Some Dynamics of Disadvantaged Students in Learning Situations." *Journal of Negro Education* 36: 315-319; 1967.
- Boocock, Sarane S. and Schild, E. O. (editors). *Simulation Games in Learning*. Beverly Hills, Calif.: Sage, 1968.
- Bradley, Nolan E. "The Negro Undergraduate Student: Factors Relative to Performance in Predominantly White State Colleges and Universities in Tennessee." *Journal of Negro Education* 6: 15-23; 1967.
- Brim, O. G. and Wheeler, S. *Socialization After Childhood: Two Essays*. New York: John Wiley, 1966.
- Bush, Dixon. "Disadvantaged Students at College: A New Dimension." *A. H. E. College and University Bulletin* 19: 2-3; 1967.
- Churchman, C. W. "Humanizing Education." *The Center Magazine* 1, No. 7: 90-93. Santa Barbara, Calif.: Center for the Study of Democratic Institutions, 1968.

- Clark, Kenneth B. and Plotkin, L. *The Negro Student at Integrated Colleges*. New York: National Scholarship Service and Fund for Negro Students, 1963.
- Creager, John A. and Astin, Alexander W. "Alternative Methods of Describing Characteristics of Colleges and Universities." *Educational and Psychological Measurement* 28: 719-734; 1968.
- Fairweather, G. W. *Methods for Experimental Social Innovation*. New York: John Wiley, 1967.
- Flacks, Richard. "The Liberated Generation: Roots of Student Protest." *Journal of Social Issues* 23: 52-75; 1967.
- Ford, D. H. and Urban, H. B. "College Dropouts: Successes or Failures?" *The College Dropout and the Utilization of Talent* (edited by L. A. Pervin, L. E. Reik, and W. Dalrymple). Princeton, N. J.: Princeton Univ., 1966.
- Freedman, M. B. *The College Experience*. San Francisco: Jossey-Bass, 1967.
- Galbraith, J. K. *The New Industrial State*. Boston: Houghton Mifflin, 1967b.
- Gales, Kathleen E. "A Campus Revolution." *The British Journal of Sociology* 17: 1-19; 1966.
- Heist, Paul. *Education for Creativity: A Modern Myth*. Berkeley: University of California, 1967. "Creative Students: College Transients," pp. 21-48.
- Holland, J. L. and Richards, J. M. "Academic and Non-Academic Accomplishment: Correlated or Uncorrelated?" *Journal of Educational Psychology* 56: 165-174; 1965.
- Hoyt, Donald P. *The Relationship Between College Grades and Adult Achievement—A Review of the Literature*. American College Testing Program Research Report No. 7. Iowa City: ACT Program, 1965.
- Hoyt, D. P. *Forecasting Academic Success in Specific Colleges*. American College Testing Program Research Report No. 27. Iowa City: ACT Program, 1968.
- Jencks, C. and Riesman, David. *The Academic Revolution*. New York: Doubleday, 1968.
- Katz, D. and Kahn, R. L. *The Social Psychology of Organizations*. New York: John Wiley, 1966.
- Katz, J. (editor). *No Time for Youth*. San Francisco: Jossey-Bass, 1968.
- Kauffman, J. F. et al. *The Student in Higher Education*. New Haven, Conn.: The Hazen Foundation, 1968.
- Kells, H. R. and Steward, C. T. "The Conference on the Cluster College Concept—A Summary of the Working Session." *Journal of Higher Education* 38: 359-363; 1967.
- Keniston, Kenneth. *Youth Radical: Notes on Committed Youth*. New York: Harcourt, Brace & World, 1968.
- Knoell, Dorothy M. "Free Choice vs. Planned Accommodation: Contrasting State Approaches to Student Input." *Research on Academic Input*. (Compiled by Clarence H. Bagley.) New York: Association for Institutional Research, 1966.
- Koile, Earl A. "Student Affairs: Forever the Bridesmaid." *NASPA Journal* 4: 65-72; 1966.
- Korn, H. A. "Counseling and Teaching: An Integrated View." *Journal of College Student Personnel* 7, 137-140; 1966.
- Korn, H. A. "Personality Scale Changes From the Freshman Year to the Senior Year." *No Time for Youth* (edited by J. Katz). San Francisco: Jossey-Bass, 1968.
- Lavin, David E. *The Prediction of Academic Performance—A Theoretical Analysis and Review of Research*. New York: Russell Sage Found., 1965.
- McClain, Edwin W. "Personality Characteristics of Negro College Students in the South—A Recent Appraisal." *Journal of Negro Education* 36: 300-325; 1967.
- McKeachie, Wilbert J. "Significant Student and Faculty Characteristics Relevant to Personality in Higher Education." *The Individual and the System* (Edited by W. J. Minter.) Ninth Annual Institute on College Self-Study. Boulder, Colo.: Western Interstate Commission for Higher Education, 1967.
- Madison, Peter. *Personality in College*. Reading, Mass.: Addison-Wesley, 1966.
- Marcuse, H. *One-Dimensional Man*. Boston: Beacon, 1964.
- Martin, Warren B. *Alternative to Irrelevance: A Strategy for Reform in Higher Education*. Nashville, Tenn.: Abingdon, 1968.
- Minter, W. J. (editor). *The Individual and the System*. Ninth Annual Institute on College Self-Study. Boulder, Colo.: Western Interstate Commission for Higher Education, 1967.
- Mumford, L. *The Myth of the Machine*. New York: Harcourt, Brace, 1966.
- Newcomb, Theodore. "The Contribution of the Interpersonal Environment to Students Learning." *NASPA Journal* 5: 175-178; 1967.

- Newcomb, T. M. and Feldman, K. A. *The Impact of Colleges Upon Their Students*. A Report to the Carnegie Foundation for the Advancement of Teaching, January 1968.
- Newcomb, T. M. and Wilson, E. K. (editors). *College Peer Groups*. Chicago: Aldine, 1966.
- Newcomb, T. M. et al. *Persistence and Change: Bennington College and Its Students After Twenty-Five Years*. New York: John Wiley, 1967.
- Nixon, James. "The Learner's View About Personalizing Higher Education." *The Individual and the System* (Edited by W. J. Minter.) Ninth Annual Institute on College Self Study. Boulder, Colo.: Western Interstate Commission for Higher Education, 1967.
- Parsons, T. and Clark, K. B. (editors). *The Negro American*. Boston: Houghton Mifflin, 1966.
- Pervin, Lawrence A. "Reality and Nonreality in Student Expectations of College." *The Journal of Psychology* 64: 41-48; 1966.
- Pervin, L. A.; Reik, L. E.; and Dalrymple, W. (editors). *The College Dropout and the Utilization of Talent*. Princeton, N. J.: Princeton Univ., 1966.
- Potter, P. "Student Discontent and Campus Reform." *Order and Freedom on the Campus*. Seventh Annual Institute on College Self Study. Boulder, Colo.: Western Interstate Commission for Higher Education, 1965.
- Richards, J. M. and Lutz, Sandra. *Predicting Student Accomplishment in College from the ACT Assessment*. Iowa City: American College Testing Program, 1967.
- Rokeach, Milton. *Beliefs, Attitudes, and Values*. San Francisco: Jossey-Bass, 1968.
- Rudolph, Frederick. "Changing Patterns of Authority and Influence." *Order and Freedom on the Campus*. Seventh Annual Institute on College Self Study. Boulder, Colo.: Western Interstate Commission for Higher Education, 1965.
- Sabine, Gordon A. "Michigan State's Search for More Negro Students." *College Board Review* 69: 11-14; 1968.
- Sampson, E. "Stirrings Out of Apathy: Student Activism and the Decade of Protest." *Journal of Social Issues* 23, No. 3; 1967.
- Sanford, Nevitt (editor). *College and Character*. (A brief version of *The American College*). New York: John Wiley, 1964.
- Sanford, Nevitt. *Where Colleges Fail*. San Francisco: Jossey-Bass, 1967.
- Stanley, Julian C. and Porter, Andrew C. "Correlations of Scholastic Aptitude Test Scores with College Grades for Negroes Versus Whites." *Journal of Educational Measurement* 4: 199-218; 1967.
- Thoresen, Carl E. "Oral Non-Participation in College Students: A Study of Characteristics." *American Educational Research Journal* 3: 198; 1966a.
- Thoresen, Carl E. "Video in the College Classroom: An Exploratory Study." *Personnel and Guidance Journal* 45: 144-149; 1966b.
- Trent, J. W. and Medsker, L. L. *Beyond High School*. San Francisco: Jossey-Bass, 1968.
- Univ. of Calif., Berkeley Academic Senate. *Education at Berkeley*. Report of the Select Committee on Education. Berkeley: Regents of the University, 1966.
- Webb, E. J. et al. *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. New York: Rand McNally, 1966.
- Williamson, E. G. and Cowan, J. L. *The American Student's Freedom of Expression*. Minneapolis: Univ. of Minn., 1966.

### Additional References

- Axelrod, J. "An Experimental College Model." *Educational Record*: 327-337; 1967.
- Axelrod, J. "The Student and the Grading System." *The Creative College Student*. (Edited by Paul Heist.) San Francisco: Jossey-Bass, 1968.
- Baratz, Stephen S. "Effect of Race of Experimenter, Instructions and Comparison Population Upon Level of Reported Anxiety in Negro Subjects." *Journal of Personality and Social Psychology* 7: 194-196; 1967.
- Barton, A. H. *Organizational Measurement and Its Bearing on the Study of College Environments*. New York: College Entrance Examination Board, 1961.
- Barton, A. H. "Studying the Effects of College Education." *The College Student and His Culture*. (Edited by K. Yamamoto.) Boston: Houghton Mifflin, 1968.



- Boocock, Sarane. "Toward a Sociology of Learning: A Selective Review of Existing Research." *Sociology of Education* 39, No. 1: 1-45; 1966.
- Bowers, J. E. "A Test of Variation in Grading Standards." *Educational and Psychological Measurement* 27: 429-430; 1967.
- Buckley, Walter (editor). *Modern Systems Research for the Behavioral Scientist*. Chicago: Aldine, 1968.
- Clark, Kenneth B. "Higher Education for Negroes: Challenges and Prospects." *Journal of Negro Education* 36: 196-203; 1967.
- Davis, Sheldon A. "An Organic Problem-Solving Method of Organizational Change." *Journal of Applied Behavioral Science* 3, No. 1: 8-20, 1967.
- Dreger, R. M. and Miller, K. S. "Comparative Psychological Studies of Negroes and Whites in the United States: 1959-65." *Psychological Bulletin Monograph Supplement* 70, No. 3: Part 2: 1-58; 1968.
- Dressel, Paul L. "Factors Involved in Changing the Values of College Students." *Educational Record* 46: 104-113; 1965a.
- Dressel, Paul L. "A Look at New Curriculum Models for Undergraduate Education." *Journal of Higher Education* 36: 89; 1965b.
- Dressel, Paul L., and Lehmann, Irvin J. "The Impact of Higher Education on Student Attitudes, Values and Critical Thinking Abilities." *Educational Record* 46: 248-259; 1965.
- Galbraith, John K. "Facing Political Reality: How the University Can Protect Itself." *College Management* 2: 32-36; September 1967a.
- Harris, C. (editor). *Problems in Measuring Change*. Madison: Univ. of Wisc., 1963.
- Harris, J. and Reitzel, John. "Negro Freshman Performance in a Predominantly Non-Negro University." *Journal of College Student Personnel* 8: 366-368; 1967.
- Harrison, E. C. "Improving Negro Colleges." *Phi Delta Kappan* 48: 296-298; 1967.
- Juola, Arvo E. "Illustrative Problems in College-Level Grading." *Personnel and Guidance Journal* 47: 29; 1968.
- Pallet, J. E. and Hoyt, D. P. "College Curriculum and Success in General Business." *Journal of College Personnel* 9: 238-245; 1968.
- Pervin, Lawrence A. "The College as a Social System." *Journal of Higher Education* 38: 317-322; 1967.
- Plant, W. T. "Longitudinal Changes in Intolerance and Authoritarianism for Subjects Differing in Amount of College Education Over Four Years." *Genetic Psychology Monograph* 72: 247-287; 1965.
- Thoresen, Carl. "The Systems Approach in Counselor Education: Basic Features and Implications." *Counselor Education and Supervision*, in press.
- Webb, Sam C. "The Relations of College Grades and Personal Qualities Considered Within Two Frames of Reference." *Multivariate Behavioral Research Monographs* No. 67: 2; 1967.

#### AUTHOR

Korn, Harold A. Address: Stanford University at Stanford, California Title: Associate Director, Counseling and Testing Center Age: 39 Degrees: B.S., C.C.N.Y.; M.A., C.C.N.Y.; Ph.D., Univ. of Minn. Specialization: Personality Assessment; Vocational Development Theory.





## 4: COUNSELOR EDUCATION

John M. Whiteley\*

Washington University

This chapter represents a broadened undertaking for the *Review of Educational Research*, reflecting the recent emergence of counselor education into a more professional status. As with any developing profession some areas of counselor education have progressed faster than others. For example, such vital areas as the evaluation of counselor education programs, the problems of professional ethics, and the counseling of children with special problems such as the ghetto child have received very little attention. In contrast, such areas as standards, the content of counselor education programs, the use of sub-professional support personnel, and the problems of counselor selection and role have received considerable attention.

### *The Selection and Role of the Counselor*

The role of the counselor has been frequently considered apart from the problems involved in the selection of counselors. Historically, the justification for this has been that as an emerging profession, counseling has needed to focus on its relationship to teaching and the educational process. Mosher (1967) observed that the issues of selection and training are inseparable from the conceptual issue of what effects are to be produced by

---

\*Dr. Garry Walz, University of Michigan, and Dr. Gail Farwell, University of Wisconsin, served as consultants to Dr. Whiteley on the preparation of this chapter.

the counselor. Mosher emphasized that unless counseling objectives are kept continually in the forefront, research on selection and effectiveness can become conceptually unrelated or unimportant. Regrettably, the bulk of the literature on problems of counselor selection has been divorced from what the counselor is to accomplish and how he is to accomplish it.

Hill and Green (1960) found no "major" longitudinal study of selection, training and evaluation. Their finding remains true today. Most counselor selection studies may be classified under four headings: 1) studies of differences between counselors and non-counselors, 2) studies of differences between effective and ineffective counselors, 3) prediction studies, and 4) trait and characteristic studies.

### *Counselors and Non-Counselors*

Although popular, the descriptive studies of the differences between counselors and some norm groups of non-counselors have been of marginal value. Scores on paper and pencil inventory tests, such as the California Personality Inventory, the Guilford-Zimmerman Temperament Survey, the Edwards Personal Preference Schedule, the Allport-Vernon-Lindzey Study of Values, and the Minnesota Multiphasic Personality Inventory, have typically been employed in these descriptions. In reviewing these studies, Patterson (1967b) pointed out that the differences found are typically so small that they are of minimal practical significance, that the scores of counseling students differ little from other college graduate students, and that effectiveness in counseling and the question of what role the counselors may be implementing are not considered.

### *Effective and Ineffective Counselors*

Another popular, if equally unproductive, approach has been to describe the differences between effective and ineffective counselors. Paper and pencil tests have often been used in these studies (see Patterson, 1967b). There are several limitations to this type of study, some of which were pointed out by Sprinthall, Whiteley and Mosher (1966): 1) intuitive "global" judgments of effectiveness are used which cannot be replicated, 2) there is no specification of what kinds of behavior are used to differentiate trainees, and 3) explicit guidelines and theoretical justifications for how students were rated are lacking. Patterson (1967b) observed that the differences found between effective and ineffective counselors were significant but too small to be important in selection. It is not clear from such studies whether the differences found existed prior to training.

### *Prediction Studies*

Prediction studies have been conducted in which researchers normally employ admission criteria as the independent variable and practicum ratings on a semester's supervised experience as the dependent variable (see Patterson, 1967b). An explicitly stated theoretical rationale is typically

lacking in these studies, which may account for failures in replication. A more serious deficiency of prediction studies stems from the researcher's inattentiveness to the complexity of the results he is attempting to predict. Does the researcher have a valid reason for assuming that the student's academic achievement in such courses as Tests and Measurements or the Psychology of Personality is in any way related to creating effective one-to-one counseling? And, given some minimal level of intelligence (whatever it is—and it undoubtedly varies among institutions), is the researcher justified in neglecting to question which personal qualities are associated with effective counseling? These problems must be studied together.

While the focus of prediction studies has been on "performance" in counseling, the studies have been performed predominantly with counselors who are still in training. This group is obviously convenient to research since normally they constitute a "captive" population. Because the profession is just beginning to learn the intricacies of predicting effectiveness, research on counselors in training is a convenient starting place. However, it should be emphasized that success in training—and its prediction—represents at best only a proximate criterion. The ultimate criterion is an effective performance as a counselor on the job.

### *Traits and Characteristics of Counselors*

The study of counselor characteristics or traits has been a traditional area for research in counselor selection. The assumption underlying this research is that certain personality traits such as sociability or friendliness are relevant dimensions within the personality organization of the counselor. However, this assumption is questionable. Efforts to evaluate personal characteristics have been sporadic and unrelated (Cottle, 1953). Research on counselor characteristics has failed generally to relate a specific rationale for particular traits to a theory of counseling—except in the most general sense. For example, a typical rationale might be that if counseling includes working with people, then the counselor ought to score high on a trait of "liking" people. Descriptions of the counselor as a person range from qualities such as "Belief in each individual" and "Commitment to individual values" (ASCA, 1965a), a "sense of mission" (Cox, 1945), integrity, vitality and judgment (Graves, 1944) to the counselor "as a woman" (Farson, 1954).

It seems fruitless to continue this inquiry of particular traits. Human qualities may indeed be relevant to counseling, but the consideration of traits per se is not going to reveal the relationship of human qualities to effective counseling. The focus must be shifted from what the counselor is in terms of a static model to an evaluation of what the counselor actually does that is effective. By assessing what the counselor does (his actual behavior), criteria of competence can be developed. These criteria may then be related to higher-order personality dimensions for use in selection.



By attempting first to specify the counselor behaviors which are associated with client change, it becomes possible to relate these behaviors to higher-order personality factors within a systematic rationale. Sprinthall, Whiteley and Mosher (1966), for example, used this approach by starting with specific counselor behaviors and relating them to cognitive flexibility. Allen and Whiteley (1968) similarly examined counselor behavior in relation to psychological "openness."

It is vital to keep in the forefront the conceptual issue of what effects the counselor is to produce. What should the counselor accomplish? How should he behave? The American School Counselors Association (1965a, 1965b) has attempted to clarify the role of the secondary school counselor. A survey of the literature on counselor role is beyond the scope of this particular review. However, it is necessary to relate the role of the counselor in particular settings to problems in counselor selection. Counselor educators must clarify what effects the counselor is to produce and specifically how he is to produce them *before* substantive progress can be made on either role or selection problems.

### *Goals of Counseling*

There is lively debate centered on the question of goals. Patterson (1964, p. 125) stated that, "The client centered counselor does have the same goal for all clients—essentially maximizing freedom of specific choices of behavior to allow maximum self-actualization." A different position was taken by Krumboltz (1966, p. 154) when he said, "the goals of counseling should be capable of being stated differently for each individual client." Blocher (1965) sharpened another area of disagreement that affects counselor educators when he declared that two distinct alternatives were open to counseling. One alternative is the existential position, which seems a "philosophically attractive" but "scientifically unclear" path. The other alternative is the behavioristic road, which is "scientifically rigorous" but "philosophically frightening." Such controversies are far from being resolved. Counselor educators continue to profess to be training counselors. At best they differentiate "counselor" into elementary counselor and high-school counselor. They avoid the basic issues of what the counselor should accomplish and how he is to accomplish it. Further specification of the desired outcomes of a training experience and how to measure these outcomes is needed. Only after such specification can the problem of selection and role be meaningfully answered.

### *The Content of Counselor Education Programs*

Coursework practicum innovations have received considerable attention. The introductory guidance course was considered by Scott (1968), Joslin (1968), Tiedeman (1968) and Farwell (1968). Their papers covered

a variety of issues, including the importance given to counseling as a guidance function, the role of philosophy, the proper emphasis on research, approaches to a self-understanding, and the relationship of the introductory guidance course to the practicum. Tolbert (1966) considered the basic guidance course in terms of whether it is really necessary, what it should contain, how it should be presented, who should take and teach it, and the research which is necessary on it.

Lister (1967) outlined the difficulties which accrue to counselor educators and subsequently to the profession from the phenomenon of "theory aversion." Theory aversion refers to the "counselor candidates' distaste for content emphasizing the principles, assumptions, objectives, philosophy or ethics which constitute the general conceptual framework for counseling and guidance practice" (p. 91). The centrality of theory to nearly all introductory counseling courses is accepted. Brammer's (1966) discussion of teaching counseling theory, the contribution that theory makes to the practice of counseling, and the necessary steps involved in developing a theory is pertinent to the problem of "theory aversion."

A number of innovative approaches to coursework appeared in the literature. Meek and Parker (1966) described a method designed to improve the efficiency of instruction and to coordinate an academic course in counseling with a course in supervised practice. Group counseling work was considered by Bonney and Gazda (1966) and Foreman (1967). Bonney and Gazda focused on the question of expecting or requiring students in counselor education programs to accept personal counseling for increased self-understanding. They reported a largely successful experiment with the use of group counseling. Foreman discussed the approach to group counseling of the National Training Laboratories. The emphases of these group experiences were on the expression of feeling, listening, and helping others to express themselves.

Innovations in practicum have focused primarily on changes in structure and on the proper balance between on-campus and off-campus experiences, with identification made of the special contributions of each. Patterson (1968) raised important ethical issues about practica by reviewing the evidence that counseling or psychotherapy, even when practiced by experienced counselors or therapists, may lead to client deterioration. He expressed his concern for the trend toward an earlier practicum, which he felt raised a substantive ethical issue. Hasely and Peters (1966) outlined two recurrent problems with practica: securing appropriate clients for the on-campus counseling experience and providing a school setting that realistically relates counseling to the total guidance program. Hansen and Moore (1966) elaborated the potentialities in off-campus practica. Boy and Pine (1966) have drawn attention to the need for coordination between the counselor education institution and the school systems providing the field-work setting. Aubrey (1967) pointed out the critical importance of the first two weeks of the practicum.

## *Approaches to Supervision*

Supervision as a topic has too frequently led to a discussion of whether counseling or teaching is at issue. Such debate has regrettably served to obscure the differences in the definitions of supervision, counseling, and teaching that are implicitly employed. Mosher (1968, p. 1-2) provided a perspective on the problem of defining supervision in teaching when he noted that "when we know more about *what* to teach, *how* and with what *specific* effects on students, we will be much less vague about supervision." When Mosher's statement is translated into the problems of defining supervision in counseling, several key questions emerge. What constitutes effective counseling? How can effective counseling be measured? What are the appropriate criteria of counseling effectiveness? Without answers to these questions, it is unlikely that current disagreements about what supervision is and how it should be done will be resolved.

Several different approaches to studying supervision have appeared in the recent literature. One line of inquiry has centered on supervisors' expectations or perceptions of their role. Walz and Roeber (1962) studied the current orientations and procedures in the supervision of counselor trainees. Gysbers and Johnston (1965) found a diversity of opinion among supervisors on what their role should be, particularly on how specific or directive they should be with trainees. In a related study, Johnston and Gysbers (1966) found that supervisors expressed a preference for democratic relationships with counselor candidates, deeming paternalistic or laissez-faire relationships inappropriate. Davidson and Emmer (1966) studied the relationship between supervisor behavior and counselor change. They found that nonsupportive behavior by supervisors causes the counselor-in-training to focus more on himself and less on the client. Further research is needed to identify how other supervisor behaviors affect counselors-in-training.

A second line of inquiry has centered upon the perceptions and expectations that counselors-in-training have of their supervisors. Delaney and Moore (1966) found that before supervision began, counselors perceived their supervisors as being primarily instructors. Gysbers and Johnston (1965) found that counselors wanted their supervisor to supply extensive help in dealing with initial counseling contacts and to provide detailed factual information about counseling techniques, test information and relevant reading assignments. Counselors expected their supervisors both to counsel them and to evaluate them. Hansen (1965) used a relationship inventory based on Rogers's necessary and sufficient conditions for personality change to assess supervisees' expectations of their supervisors. Miller and Oetting (1966) studied the characteristics of supervision which counselors-in-training identified as important.

The importance of relating the change in supervisee behavior to the approaches to supervision has been discussed earlier. One of the most

promising approaches to this problem was developed by Ivey et al. (1968) in their studies of "micro-counseling." Micro-counseling refers to a method of video instruction of counselors in specific base skills. The authors have developed micro-counseling training procedures for three different skills: attending behavior, reflection of feeling, and summarization of feeling. Observing that beginning counselors can learn these skills readily in a short time period, they (Ivey et al., 1968, p. 3) commented, "... attending behavior and its related concepts of reflection and summarization of feeling may be described in behavioral terms meaningful to beginning counselors." Although the micro-counseling method is still in the early stages of its own development, it appears to be particularly promising for the development of specific counselor skills. Further research is necessary to determine what other skills may be imparted through this method, and how micro-counseling compares to other supervisory techniques in developing the same skills.

### *Evaluating the Effects of Counselor Education*

Despite the importance of evaluating counselor education and its centrality to the profession, very little research on evaluation has been conducted. The few evaluations that have been done were generally of low quality, superficial and so narrowly defined that misleading implications could be drawn. Regrettably, evaluation does not appear to be a term with any substance in counselor education programs.

Why is there so little good evaluation of counselor education? Arnold (1962, p. 189) said that counselor educators "simply do not know what they are doing, nor how to evaluate it." Meyering (1964, p. 37) offered his opinion: "Counselor educators are basically an uncreative lot. For the most part, we are intellectually lazy, inefficient, egocentric, and have a real commitment to maintenance of the status quo." Counselors have done little to disprove these allegations, if judged by the quality and frequency of their evaluations.

The current stage of professional development in counseling suggests one reason for the lack of good evaluation. The American School Counselor Association (1965a) has only recently developed a role statement for secondary school counselors. Even more recently, the Association for Counselor Education and Supervision (1967a) completed their revised *Standards*. These documents should provide initial guidelines for a consideration of what is to be accomplished in an evaluation of the effects of counselor education. However, these documents do not provide sufficient detail about the performance criteria by which to conduct a program evaluation that will accurately assess counselor performance and client change. An outline of what a "good" program should provide is a very limited type of "evaluation."



Another reason for the lack of evaluation is that counselor education is affected by the same issues which continue to plague research on the selection of counselors. It will not be possible to study with any precision the effects of a training program until it becomes clearer what effects the counselor should produce and how he is to accomplish his purposes. When the desired outcomes of training programs are further specified and more is known about what counselors do that makes their training worthwhile, evaluation will become a more valid and useful undertaking.

The evaluation studies which have been completed are limited in scope. Demos and Zuwaylif (1962) studied the changes in counselor attitudes in relation to the theoretical positions of their supervisors. Attitude change was also the subject of previous studies by Hopke (1955); Demos and Zuwaylif (1962); Jones (1963); Webb and Harris (1963); Wrightsman, Richard and Noble (1966); and Munger and Johnson (1960). These studies, however, did not assess what effects in actual counseling behavior occurred from attitude change on the part of trainees.

Peters and Thompson (1968) surveyed how school superintendents view counselor preparation. Shertzer and England (1968) studied counselor opinions about their preparation in a questionnaire survey. Graduates of their counseling program were asked for descriptive information about their current employment, preparation, satisfaction, and important training experiences.

Schoch (1966) studied the effects of a summer counseling practicum on counselor behavior in a counseling relationship. Using an instrument developed by Combs and Soper (1963) to define "good" counselors, Schoch found in essence that the counselors became "better" on the relevant dimensions as a result of the practicum experience. Schoch raised a number of pertinent questions, unanswered by his study, about the effect on the finding of the stratified nature of his sample, the shortness of the program, the permanence (or lack) of the changes, the need for a follow-up study and the need for replication with other groups.

Seamen and Wurtz (1968) reported a study on evaluating the practicum. They approached this critical question by administering a test of sensitivity before and after the practicum experience. The dependent variables consisted of peer and staff ratings. The deficiencies of this study illustrate the research problems which continue to handicap investigations in evaluating counseling programs. The sensitivity test used by Seamen and Wurtz was the "Experimental Test of Counselor Sensitivity" developed by Hood and reported on by Sundberg (1952). There are no further references to the characteristics of the test. It *may* be a good test for the purpose, but it is relatively unknown and the Sundberg reference is an unpublished doctoral dissertation. The sample size was only nine students. The dependent variables were so global that replication is impossible. To obtain the peer ratings, at the last practicum session "all students

were asked to rank all members of the group, including themselves, in terms of perceived potential effectiveness as a counselor" (p. 282). No information is provided about what they considered to be "perceived potential effectiveness." The criteria used by the three supervisors is no more specific. The point in reviewing this study was to illustrate several important areas of deficiency in conducting and reporting research concerned with evaluating the effects of counselor training programs. Seaman and Wurtz had some good ideas, but the design of their study and the manner in which it was reported severely handicap the usefulness of their research.

Thoresen (1968) offered a paper on the relevance of the systems approach to counselor education in specifying objectives, conceptualizing problems, gathering information and evaluating solutions. He emphasized the futility of attempting to evaluate counselor education programs without clearly stated objectives which are in some way measurable; he also stressed the need for giving attention to the entire program, not just to one part of it.

A differentiated approach to counselor preparation has been proposed by Krumboltz and Thoresen (1968); this approach has important implications for evaluation. They called for the training of a "new breed of specialists" who would recognize that no current approach constitutes the sole answer. Thoresen and Krumboltz advocate using a systems approach (Thoresen, 1968) and integrating all existing approaches within an empirically functioning system that would create evidence of what works. Further, they proposed to eliminate the configuration of courses and practicum experiences as they are now conceptualized and structured.

The following questions would be central to the Thoresen and Krumboltz systems approach: What are the performance goals of the program? What experiences will "best" (in terms of efficiency of resources and effectiveness) produce these performances? What are the possible interactive factors (trainee characteristics, initial skills, certain client problems, particular environmental settings) which lead to differential training programs? They emphasize that no one training package would be envisaged for all trainees, but that evaluation would be related to the effectiveness of a variety of empirically validated experiences for different trainees. This approach should effectively progress toward eliminating the situation that enabled Carkhuff et al. (1968) to cite the negative effects of traditional training programs.

## *Special Issues in Counselor Education*

### *Support Personnel*

The topic of support personnel has been a concern of APGA which, through its committee on professional preparation and standards, issued a statement (APGA, 1967) on the preparation for technical and non-technical

roles. However, Munger (1968), after participating in five regional ACES meetings, found that counselor educators "expressed little interest" in support personnel. Gust (1968) raised some important issues surrounding the role of support personnel. Salim and Vogan (1968) outlined several functions which they felt support personnel could perform effectively.

There is an increasing body of literature concerned with the use of lay or minimally trained personnel in personal counseling (Rioch et al., 1963; Poser, 1966; Beck, Kantor, and Gelineau, 1963; Carkhuff and Truax, 1965a, 1965b; Truax and Carkhuff, 1967; Carkhuff, 1966; and Golann and Magoon, 1966). Golann and Magoon (1966), using carefully selected individuals who did not have professional degrees, found that these individuals could be trained to provide psychotherapeutic services in school settings. Carkhuff (1966), summarizing the literature, concluded that "the primary conditions of effective treatment are conditions which minimally trained non-professional persons can provide."

### *Standards and Accrediting in Counselor Education*

In 1964 the Association for Counselor Education and Supervision (ACES, 1964) adopted a set of *Standards* recommended by their committee; these standards were intended for voluntary use on a three-year experimental basis. After the trial period, the suggested changes were incorporated into a set of revised *Standards*. In 1967 the revised *Standards* were adopted by ACES (Hill, 1967a; ACES, 1967a) for use in counselor education. A manual (ACES, 1967b) was designed to assist in the use of the *Standards* and to provide suggestions for institutional self-evaluations. Mazer (1967) offered an analysis of the possible reasons for what he considered to be poor participation. Hill (1967b) said that he had been encouraged by the extent of participation in the *Standards* project. Hill felt that participation had been higher than usual in professional association matters. Patterson (1967a) took sharp issue with the thirty-hour practicum requirement for first-year counselor trainees; he felt that the statement was too specific for standards which were likely to become the basis for accreditation. Hill (1967a) provided a highly useful perspective on the *Standards* and on the developments based upon the *Standards* which are needed.

Although progress in the development of standards for counselor education has been a significant accomplishment of the profession in the past several years, accreditation procedures are far from satisfactory at this time. An important task for the immediate future is the establishment of a more orderly and qualitatively sound system of accreditation.

The overwhelming majority of counselor education programs are now offered within the education faculties of colleges and universities. Consequently, many counselor education programs are under the purview of the National Council for Accreditation of Teacher Education (NCATE).

The situation with NCATE, according to Stripling (1968, p. 201) is that "no specific criteria for accrediting counselor education have been developed; and, in many cases, no qualified supervisor, counselor, or counselor educator, is on an institutional visiting committee." Stripling (1968, p. 202) reported that NCATE has "expressed a willingness to assist APGA in the development of working relationships which would lead to a more satisfactory accrediting procedure." Although nominal progress toward developing working relationships has been made, much needs to be done. Since accreditation is highly important to the ultimate quality of professional service, the relationship with NCATE must be assigned a high priority.

### *The Future of Counselor Education*

This review of the research in counselor education has demonstrated that over the past three years there have been areas of notable accomplishment. But it is also strikingly clear that there is only the barest knowledge about some of the central issues in counseling—particularly as they relate to problems in counselor education.

If major progress is to occur, there must be a great increase in learning about the nature of the counseling process, specifically learning what counselors do that is helpful to particular clients with different problems. Recent theoretical advances in client-centered counseling and behavioral counseling have been supported by careful empirical research. However, more research by proponents of both approaches is necessary and is being conducted. Some of the remaining major problems concern the goals of the counseling endeavor, the criteria by which those goals are measured, and the assessment of the extent to which goals are reached with particular clients.

Regarding counselor education, there is the additional problem of translating the knowledge to be gained about counseling into the teaching process. Formerly, counselor educators (because of theoretical confusions in counseling) have avoided the central questions about the effects the counselor is to produce and what he can do to accomplish his goals. The results of this avoidance have been to add to the confusion of counselors entering the field, to restrict the structure of counselor education programs, to create teaching systems which have no feedback mechanisms, and to make evaluating the effects of counselor education programs more art than science.

Perhaps the most pressing immediate needs are to recognize that the service role of the counselor is paramount and that counselor training must be organized to help counselors learn and use alternate ways of assisting clients to fulfill their goals. There is also a need to refocus the selection of counselors in terms of the questions concerning what effects the counselor



is to produce. In addition, further specification of the desired outcomes of a training experience and how to demonstrate those outcomes merit careful attention. Progress has been made recently in developing specific counselor skills through different training experiences. Further research is needed to determine the range of specific skills that can be taught by such approaches as micro-counseling. In addition, researchers must determine which skills cannot be taught through micro-counseling and which other methods of supervision might be appropriate for teaching them.

Evaluating the effects of counselor education programs has received scant attention. As the objectives of total training programs are explicated, it will be possible to evaluate them in terms of their goals. Alternate approaches to the goals can be devised and their effectiveness measured. Achieving adequate program evaluation will take considerable time. A set of minimal standards in counselor training has only recently been established. As accreditation is currently practiced, it lacks substance. The state of understanding of the nature of counseling makes impossible an early closure on what should be included in a training program.

The immediate future, if it is not to be a repetition of the past, must include at least the following: 1) closer attention to the results of research on the counseling process in formulating curricula, 2) careful study of what supervisors can do to enhance skill development in counselors-in-training, 3) establishment of objectives which can be measured for counselor training programs as well as a commitment to evaluate graduates in terms of how they meet those objectives, and 4) the construction of a systematic method for modifying counselor education programs in terms of those evaluations.

### Bibliography

- Allen, Thomas W. and Whiteley, John M. *Dimensions of Effective Counseling*. Columbus: Charles E. Merrill, 1968.
- American Personnel and Guidance Association. "Support Personnel for the Counselor: Their Technical and Non-Technical Roles and Preparation." *Personnel and Guidance Journal* 45: 857-861; 1967.
- American School Counselor Association. "Statement of Policy for Secondary School Counselors." *Counseling: A Growing Profession*. (Edited by John W. Loughary, Robert O. Stripling, and Paul W. Fitzgerald.) Washington, D. C.: Amer. Personnel and Guidance Assoc., 1965a. Pp. 94-99.
- American School Counselor Association. "Guidelines for Implementation of the ASCA Statement of Policy for Secondary School Counselors." *Counseling: A Growing Profession*. (Edited by John W. Loughary, Robert O. Stripling, and Paul W. Fitzgerald.) Washington, D. C.: Amer. Personnel and Guidance Assoc., 1965b. Pp. 100-106.
- Arnold, Dwight L. "Counselor Education as Responsible Self-Development." *Counselor Education and Supervision* 1: 185-192; 1962.
- Association for Counselor Education and Supervision. "Standards for Counselor Education in the Preparation of Secondary School Counselors." *Personnel and Guidance Journal* 42: 1060-1073; 1964.
- Association for Counselor Education and Supervision. "Standards for the Preparation of Secondary School Counselors." *Personnel and Guidance Journal* 46: 96-106; 1967a.
- Association for Counselor Education and Supervision. *Manual for Self-Study by a*

- Counselor Education Staff. Washington, D. C.: Amer. Personnel and Guidance Assoc., 1967b.
- Aubrey, Roger F. "Elementary School Counseling Practicum: Suggestions for Experiences and Expectations." *Counselor Education and Supervision* 7: 13-19; 1967.
- Beck, J. C.; Kantor, D.; and Gelineau, V. "Follow-up Study of Chronic Psychotic Patients "Treated" by College Case-aid Volunteers." *American Journal of Psychiatry* 120: 269-271; 1963.
- Blocher, Donald. "Issues in Counseling: Elusive and Illusional." *Personnel and Guidance Journal* 43: 796-800; 1965.
- Bonney, Warren C. and Gazda, George. "Group Counseling Experiences: Reactions by Counselor Candidates." *Counselor Education and Supervision* 5: 205-211; 1966.
- Boy, Angelo V. and Pine, Gerald T. "Strengthening the Off-Campus Practicum." *Counselor Education and Supervision* 6: 40-43; 1966.
- Brammer, Lawrence M. "Teaching Counseling Theory: Some Issues and Points of View." *Counselor Education and Supervision* 5: 120-130; 1966.
- Carkhuff, Robert R. "Training in the Counseling and Psychotherapeutic Practices: Requiem or Reveille." *Journal of Counseling Psychology* 13: 360-367; 1966.
- Carkhuff, Robert R. and Truax, Charles B. "Lay Mental Health Counseling: The Effects of Lay Group Counseling." *Journal of Consulting Psychology* 29: 426-431; 1965a.
- Carkhuff, Robert R. and Truax, Charles B. "Training in Counseling and Psychotherapy: An Evaluation of an Integrated Didactic and Experimental Approach." *Journal of Consulting Psychology* 29: 333-336; 1965b.
- Combs, Arthur and Soper, D. W. "The Perceptual Organization of Effective Counselors." *Journal of Counseling Psychology* 10: 222-226; 1963.
- Cottle, William C. "Personal Characteristics of Counselors." *Personnel and Guidance Journal* 31: 445-450; 1953.
- Cox, R. D. *Counselors and Their Work*. Harrisburg: Archive Press, 1945.
- Davidson, Terrence N. and Emmer, Edmund T. "Immediate Effect of Supportive and Nonsupportive Supervisor Behavior on Counselor Candidates' Focus of Concern." *Counselor Education and Supervision* 6: 27-31; 1966.
- Delaney, Daniel J. and Moore, James C. "Student Expectations of the Role of Practicum Supervisor." *Counselor Education and Supervision* 6: 11-17; 1966.
- Demos, George D. and Zuwaylif, Fadil H. "Counselor Attitudes in Relation to the Theoretical Positions of their Supervisors." *Counselor Education and Supervision* 2: 8-13; 1962.
- Farson, Richard E. "The Counselor Is a Woman." *Journal of Counseling Psychology* 1: 221-223; 1954.
- Farwell, Gail F. "The Introductory Course: A Synthesis of Theory and Practice." *Counselor Education and Supervision* 7: 98-101; 1968.
- Foreman, Milton E. "T Groups: Their Implications for Counselor Supervision and Preparation." *Counselor Education and Supervision* 7: 48-53; 1967.
- Golann, S. E. and Magoon, T. M. "A Non-Traditionally Trained Mental Health Counselor's Work in a School Counseling Service." *School Counselor* 14: 81-85; 1966.
- Graves, John T. "The Employee Counselor." *Occupations* 22: 495-497; 1944.
- Gust, Tim. "Support Personnel vs. The Counselor." *Counselor Education and Supervision* 7: 152-154; 1968.
- Gysbers, Norman C. and Johnston, Joseph A. "Expectations of a Practicum Supervisor's Role." *Counselor Education and Supervision* 4: 68-74; 1965.
- Hansen, James C. "Trainees' Expectations of Supervision in the Counseling Practicum." *Counselor Education and Supervision* 4: 75-80; 1965.
- Hansen, James C. and Moore, Gilbert D. "The Off-Campus Practicum." *Counselor Education and Supervision* 6: 32-39; 1966.
- Haseley, Luther L. and Peters, Herman J. "Practicum: On-Campus and Off-Campus." *Counselor Education and Supervision* 5: 141-147; 1966.
- Hill, George E. "The Profession and Standards for Counselor Education." *Counselor Education and Supervision* 6: 130-136; 1967a.
- Hill, George E. "Comment on Mazer's Comment in the Spring 1967 Issue." *Counselor Education and Supervision* 6: 353; 1967b.
- Hill, George E. and Green, Donald A. "Selection, Preparation and Professionalization of Guidance and Personnel Workers." *Review of Educational Research* 30: 115-130; 1960.

- Hopke, William E. "The Measurement of Counselor Attitudes." *Journal of Counseling Psychology* 2: 212-216; 1955.
- Ivey, Allen et al. "Microcounseling and Attending Behavior." Monograph Supplement *Journal of Counseling Psychology* 15 (Supplement): 1-12; 1968.
- Johnston, Joseph A. and Gysbers, Norman C. "Practicum Supervisory Relationships: A Majority Report." *Counselor Education and Supervision* 6: 3-10; 1966.
- Jones, Vernon. "Attitude Change in an NDEA Institute." *Personnel and Guidance Journal* 42: 387-392; 1963.
- Krumboltz, John D. "Behavioral Goals for Counseling." *Journal of Counseling Psychology* 13: 153-159; 1966.
- Krumboltz, John D. and Thoresen, Carl E. *Behavioral Systems Training Program for Counselors*. Stanford, Calif.: School of Education, Stanford Univ., May 1968. (Mimeo.)
- Lister, James L. "Theory Aversion in Counselor Education." *Counselor Education and Supervision* 6: 91-96; 1967.
- Mazer, Gilbert E. "Comment and Speculation on Hill's Report." *Counselor Education and Supervision* 6: 210; 1967.
- Meek, Clinton R. and Parker, Aileen W. "Introductory Counseling Course: Use of Practicum Students." *Counselor Education and Supervision* 5: 154-158; 1966.
- Meyering, Ralph A. "The Wonderland of Counselor Education." *Counselor Education and Supervision* 4: 37-41; 1964.
- Miller, C. Dean and Oetting, Eugene R. "Students React to Supervision." *Counselor Education and Supervision* 6: 73-74; 1966.
- Mosher, Ralph L. "Clinical Supervision: The Analysis of Teaching." *Supervision: New Dimensions*. (Edited by Ralph Mosher, David Purpel, and Kiyo Morimoto.) Boston: Houghton-Mifflin, 1968. Pp. 1-2.
- Munger, Paul F. Support Personnel. *Counselor Education and Supervision* 7: 81-82; 1968.
- Munger, Paul F. and Johnson, Carleton A. "Changes in Attitudes Associated with an NDEA Counseling and Guidance Institute." *Personnel and Guidance Journal* 38: 751-753; 1960.
- Patterson, Cecil H. "Comment." *Personnel and Guidance Journal* 43: 124-126; 1964.
- Patterson, Cecil H. The Practicum Requirement in "Standards—1967 Edition." *Counselor Education and Supervision* 6: 352-353; 1967a.
- Patterson, Cecil H. "The Selection of Counselors." *Research in Counseling*. (Edited by John M. Whiteley.) Columbus: Charles E. Merrill, 1967b.
- Patterson, Cecil H. "The Counseling Practicum: An Ethical Issue." *Counselor Education and Supervision* 7: 322-324; 1968.
- Peters, Herman J. and Thompson, Charles L. "School Superintendents View Counselor Preparation." *Counselor Education and Supervision* 7: 379-386; 1968.
- Poser, E. G. "The Effect of Therapists' Training on Group Therapeutic Outcome." *Journal of Consulting Psychology* 30: 283-289; 1966.
- Rioch, M. J. et al. "National Institute of Mental Health Pilot Study in Training Mental Health Counselors." *American Journal of Orthopsychiatry* 33: 678-689; 1963.
- Salim, Mitchell and Vogan, H. Jayne. "Selection, Training and Functions of Support Personnel in Guidance: The Counselor Assistant Project." *Counselor Education and Supervision* (Special Issue) 7: 227-236; 1968.
- Schoch, Eugene W. "Practicum Counselor's Behavioral Changes." *Counselor Education and Supervision* 6: 57-62; 1966.
- Scott, C. Winfield. "The Introductory Guidance Course: Philosophy or Survey of Practices." *Counselor Education and Supervision* 7: 83-85; 1968.
- Seamen, Earl H. and Wurtz, Robert E. "Evaluating the Practicum: Whither and Wither?" *Counselor Education and Supervision* 7: 282-285; 1968.
- Shertzer, Bruce and England, Joan. "Follow-Up Data on Counselor Education Graduates: Relevant, Self-Revealing, or What? *Counselor Education and Supervision* 7: 363-370; 1968.
- Sprinthall, Norman A.; Whiteley, John M.; and Mosher, Ralph L. "Cognitive Flexibility: A Focus for Research on Counselor Effectiveness." *Counselor Education and Supervision* 5: 188-197; 1966.
- Stripling, Robert O. "Current and Future Status of Accrediting Counselor Education." *Counseling Education and Supervision* (Special Issue) 7: 200-209; 1968.

- Sundberg, N. D. *The Relationship of Psychotherapeutic Skill and Experience to Knowledge of Other People*. Doctor's thesis. Minneapolis: Univ. of Minn., 1952. Abstract: *Dissertation Abstracts* 12: 390-391; No. 4, 1952.
- Thoresen, Carl E. "The Systems Approach and Counselor Education: Basic Features and Implications." *Counselor Education and Supervision*, in press.
- Tiedeman, David V. "The Introduction of Students and Guidance: Freedom and Responsibility through a Tale of the Social Graces." *Counselor Education and Supervision* 7: 86-92; 1968.
- Tolbert, E. L. "Basic Guidance Course: Forgotten But Not Gone." *Counselor Education and Supervision* 5: 131-139; 1966.
- Walz, Garry R. and Roeber, Edward C. "Supervisors' Reactions to a Counseling Interview." *Counselor Education and Supervision* 2: 2-7; 1962.
- Webb, Allen P. and Harris, John T. "A Semantic Differential Study of Counselors in an NDEA Institute." *Personnel and Guidance Journal* 42: 260-263; 1963.
- Wrightsmann, Lawrence; Richard, Wayne; and Noble, Frank. "Attitude Changes of Guidance Institute Participants." *Counselor Education and Supervision* 5: 212-220; 1966.

#### AUTHOR

Whiteley, John M. Address: Washington University, St. Louis, Missouri Title: Assistant Vice Chancellor and Chairman, Program in Counseling Age: 28 Degrees: A.B., Stanford Univ.; M.A., Harvard Univ.; Ed.D., Harvard Univ. Specialization: Designs in Counseling and Problems in Selecting Counselors.





## 5: CHANGES THROUGH COUNSELING

Ray E. Hosford\*

*University of Wisconsin*

Alan S. Briskin

*University of Wisconsin*

Research resulting in definitive knowledge of counseling outcomes is still wanting. Unfortunately, great separations among theoretical rationales, research, outcome criteria and practices remain. Laboratory-type research does not always find workable procedures for practitioners. Field research, conversely, often leaves the reader ignorant of both theoretical assumptions and specific results. It is almost impossible to determine whether the non-specific results support or challenge the loosely defined and executed procedures of the typical field study.

The failure to describe the nature of the treatment adequately makes many studies of little practical value. Too often the independent variable is labeled "counseling" and the dependent variable (the objective of counseling) is some personality test score. This situation is not unlike the medical patient who receives "doctoring" for severe abdominal pains, but whose cure is measured by a test on attitudes toward medicine. The knowledge that the patient was given "doctoring" and has perhaps changed his attitude toward medicine does little to help other physicians assist patients with similar problems.

A few major studies (e.g., Rothney, 1958; Volsky et al., 1965) have attempted to evaluate total counseling services. Although well-designed,

---

\*The authors were assisted by Jerry L. Davis.

such studies have most often been left holding the null hypothesis. The lack of specific outcome variables has been a major problem encountered in these and other similar studies. In retrospect, Volsky et al. (1965, p. 178) suggest "... a promising approach to criterion variables for assessing counseling outcomes is found in specific variables—variables which have certain face validity stemming from their direct reflection of client concerns in seeking and in experiencing the treatment process." Kagan's (1966) observations on confusions in group counseling research are relevant; he noted that the use of inadequate labels and the failure to describe the actual nature of treatments have retarded research in this area.

The studies reported during 1965-1968 which, in the reviewers' opinion, contributed the most toward bridging the persistent gap between research and practice have been included in this chapter. The outcome topics which were selected were those which the chapter authors found to be most prevalent among the studies dealing with clients in educational settings. Whether the counseling treatment might be viewed by some readers as psychotherapy or counseling or teaching was not an issue that influenced which articles were chosen for review. Counseling was viewed as any ethical procedure used by the counselor to help promote change in the client. In general, the investigations which were included in this chapter focused on changes in the client's attitude or behavior outside of the counseling interview. Process studies (counselor-client interaction, Kiesler, 1966) were incorporated when the in-counseling behavior was directly related to outcomes, and when the findings appeared to have significant implications for future research to promote changes through counseling.

### *Changes in Academic Achievement*

Improvement in academic functioning is one of the main goals sought by clients in educational settings (Volsky et al., 1965). There were numerous studies in which academic performance was the desired outcome. Measuring counseling effectiveness in this area remains a problem. GPA is the most often used measure, but it is viewed either as arbitrary and at best tangentially appropriate by some, (e.g., Hill, 1966) or as an excellent measure of counseling outcomes by others (e.g., Paul, 1968; Volsky et al., 1965).

#### *Elementary School*

Few studies have been concerned with promoting change in academic performance among elementary students; notable exceptions were the studies of Winkler et al. (1965) and Jackson (1966). Winkler et al. (1965) attempted to produce changes in the GPAs of 121 underachieving fourth-grade students in 22 classrooms. The subjects were randomly assigned by

classroom to one of five treatments such as: 1) individual client-centered counseling concerned with role playing school and family situations, drawing and telling stories about pictures, and doing manual tasks such as building small objects, 2) group counseling utilizing basically the same techniques as individual counseling and 3) group reading instruction using Science Research Associates Reading Laboratory Power Builders. Treatments did not promote significant differences in GPAs. Dependent variables, grades and scores on the California Test of Personality, served as measures of counseling success (objectives), but apparently received little emphasis during the 14 treatment sessions.

Jackson (1966) studied the effects of a three-year treatment on the academic development of 190 children who were identified as under-achievers. During the treatment period counselors worked with the students, parents and teachers of the experimental group. Seventh-grade marks in English, mathematics, social studies and science were among the major criteria assessed as the dependent variables. The grades of the experimental subjects were consistently better than those of the control group on 10 of the 12 criteria used. Although investigators encouraged the cooperating districts to limit their services given to the control subjects, this was impossible because controls were referred to special reading teachers, tutors, social workers, and other pupil services personnel. This fact may have attenuated the differences found between the two groups.

Whether using secondary rather than elementary-trained counselors affected outcomes was questioned in the Winkler et al. (1965) study. This question has heuristic value. Little evidence exists at present to substantiate different training for counselors employed at different educational levels. An exception is a study by Kranzler et al. (1966) which demonstrated that counselors trained to work with elementary school children were effective in helping students to raise their sociometric status in the classroom. The implication, however, that elementary-trained counselors are more effective than secondary-trained counselors in working with elementary school children cannot be made until experimental studies are carried out which compare both elementary- and secondary-trained counselors utilizing identical treatments to promote the same outcomes.

*Additional References:* Laxer et al. (1966); Stormer (1967).

### *Secondary School*

A study by McGowan (1968) is interesting because of its simplicity. McGowan examined the results of counseling 18 low-ability and low-achieving eighth-, ninth- and tenth-grade males who were matched on IQ and grade level and randomly assigned to control or experimental groups. Counseling consisted of ten-minute interviews held twice a week for the academic year; during each interview the client described his progress since the last interview. Significant differences were found in favor of the



experimental group for academic performance, attendance in school, and frequency of client initiated counselor appointments. Unfortunately, descriptive detail of what actually occurred during counseling was not given. A basic flaw in this study was the failure to control for the effect of providing non-specific attention and concern to the experimental group.

Gilliland (1968) demonstrated that small-group counseling of black high-school students for one hour per week during an academic year was more effective than no counseling in raising the students' academic achievement level. The two experimental groups, composed of seven boys and seven girls respectively, showed significantly higher vocabulary, reading and English usage scores than the eight boys and eight girls assigned to control groups.

Bates (1968) compared "traditional group counseling" (13 weekly class period meetings) and "accelerated interaction" (equal amount of time concentrated into two-day continuous session during school). Gain scores for eight measures such as GPA, attendance and effort grades were used. Subjects were matched for sex, grade, academic potential, socio-economic level and academic achievement. The traditionally counseled groups demonstrated significant improvement in GPA, attendance and effort grades when they were compared to appropriate control groups observed for the same time periods. The accelerated interaction groups demonstrated significant improvement only in increasing students' receptivity to the learning process through reducing tension and hostility. In general, the results indicated that the two group formats were not interchangeable in the public school setting. The study was innovative in using accelerated interaction groups in the schools.

Bunda and Mezzano (1968) demonstrated that a program of work experience significantly modified the academic and attendance performances of potential dropouts. Thirty high-school students identified as potential dropouts were placed on jobs for half of the school day. The subjects received "intensive" counseling, had their course load reduced and were given on-the-job supervision. Pre-post analyses showed significant gains in frequency of student attendance and in GPA. There was no indication of what constituted "intensive" counseling nor was it possible to discern what contributed to change, i.e., the work-study program, the counseling, being away from school for half the time, or a combination of these.

Benson and Blocher (1967) evaluated the effects of 18 weekly group counseling sessions on high-school students who were designated as low achievers with poor attitudes toward school. Teachers nominated 28 students who were then randomly assigned to experimental ( $N = 12$ ), control ( $N = 12$ ) and reserve groups. The counseled group showed a significant mean gain (+.51) in GPA over two semesters while control subjects experienced a reduction in mean GPA. It may be that group counseling for some students served to help them "hold their own" rather than

decline. The lack of knowledge about individual changes presents a problem which this study illustrates: counseling for the most part seeks to change the behavior of an individual rather than a group. An average gain of .51 for 12 students masks subtle individual changes; it is more important to know which subjects gained, remained the same or deteriorated.

### *College Students*

Ewing and Gilbert (1967) tested the effects of counseling on subsequent academic performance of a select group of superior college freshmen. Each student received a written invitation to participate in a counseling project. Comparisons were made among four groups: 1) counseled students ( $N = 90$ ), 2) invited but failed to take advantage of the counseling ( $N = 28$ ), 3) completed a research questionnaire but not invited for counseling ( $N = 117$ ), and 4) failed to fill out the questionnaire and neither invited nor counseled ( $N = 20$ ). Counseling was generally focused on assisting students to achieve better grades. Control subjects were specifically asked not to seek counseling unless they felt it was urgent to do so. Counseled students received significantly higher GPAs than did the non-cooperative control group (Group 4) and the invited but not counseled group (Group 2). However, no significant difference in GPA was found for the counseled group and subjects who completed the research questionnaire (Group 3). The study raises the question of whether attitudes and expectancies toward counseling rather than counseling per se is the crucial determinant of client change. Confounding of client expectancies and attitudes toward counseling resulting from past experiences has operated to obscure what changes counseling itself causes. Studies are needed which control for this client characteristic. Other studies need to determine how counselors can alter client attitudes and set toward counseling.

Roth et al. (1967) demonstrated that group counseling for two one-hour sessions per week for a semester was effective in helping 52 failing college students (freshmen, sophomores, juniors and seniors) improve their academic achievement. In the first five to eight sessions, group members identified particular study problems, e.g., daydreaming, and received prescriptions for change from the counselor. When the study difficulties were under control, the counselor used what was termed client-centered therapy to explore, for the balance of the semester, the assumed basic reasons for the student's failure; these reasons included problems of defensiveness, relationships with parents and independent decision-making. Counseled students showed a significant increase in GPA and control students indicated a chance increase. For twenty of the thirty counseled students who were still available for follow-up, the significant improvement held for one semester after termination of counseling.

A problem of the Roth study was that the control and experimental subjects could not be randomly selected from the same population. The

experimentals were randomly drawn from a group of 174 probationary students who were invited and counseled during the six-semester evaluation period. The controls were also probationary students, but they were not invited to participate. The controls were matched for age, major, and length of time in school, but were not matched for the criterion variable—GPA. The experimental group had a mean grade point of .94 at the beginning of the study, whereas the control group had 1.33. In terms of random error effect and regression toward the mean, it is not surprising that the experimentals showed a significant mean increase. When working with any skewed distribution, it is crucial that the comparison control groups be randomly drawn from the same population. It was premature for the authors to attribute the group counseling effectiveness to a therapeutic intervention technique of clarifying the dynamics of underachievement and change. The purpose of this approach, which was to eliminate defenses of poor study habits so more significant issues could be handled, was not supported. Experimental studies comparing underachievers experiencing group procedures with and without this intervention technique are needed before such conclusions can be made.

Ryan (1967) tested the use of reinforcement counseling in a residence hall setting to improve study habits, attitudes toward study and academic achievement. The study population consisted of 928 male and female students. Four treatments were employed: 1) specific cue-reinforcement in which the counselor gave specific cues, such as "What do you think about planning time for relaxation and rest?", and then verbally reinforced those client responses which suggested "good" study behavior; 2) general cue-reinforcement in which the counselor used the same procedure as specific cue-reinforcement to reinforce responses to general questions such as "What do you think about planning your schedule?"; 3) placebo counseling in which they discussed effective study techniques; and 4) inactive control in which subjects took criterion tests. Subjects in reinforcement groups demonstrated significantly higher study habit inventory scores than subjects in placebo and control groups—regardless of counselors or sex of subjects. No differences were found between the specific cue-reinforcement and the general cue-reinforcement groups. A six-month follow-up showed that reinforcement groups' study habits inventory scores increased over time; the placebo group maintained the same level and the inactive control group decreased. Changes in attitudes toward study were not found. Subjects in the reinforcement groups had a higher adjusted mean GPA following treatment than subjects in the placebo and inactive groups. An interesting aspect of the study was that a significant correlation was not found between counselee's perception of counselor effectiveness and counseling outcome as measured by the various criterion tests.

Thelen and Harris (1968) provided group counseling for 127 under-achieving college students. Three groups were identified: those who did



not respond to written invitations, those who offered to take tests only, and those who agreed to tests and the therapy sessions. The latter group was divided into a control group ( $N = 15$ ) and four experimental groups ( $N = 22$ ) which met one and one-half hours weekly for an average of 7.7 times. No description of treatment was included. The therapy groups experienced significantly greater GPA change than the control group. Subjects who took the battery of tests but declined counseling made almost as great a gain as the counseled group did. Since this group showed the greatest maladjustment according to scores on Cattell's 16 Personality Factors Questionnaire and self-rating scales, the authors suggested that the improvement in these subjects may have been due, in part, to motivation based on anxiety and self-doubt. This finding offers interesting possibilities for further research, e.g., the relationship between personality variables and various counseling treatments aimed at improving academic performance.

*Additional References:* Brown (1965); Dickenson and Truax (1966); LeMay and Christensen (1968).

### *Decision Making*

Helping clients develop decision-making skills as a primary counseling outcome has been supported frequently and for a long time in the literature (Ginzberg, 1952; McDaniel, 1956; Rothney, 1958). During the past three years considerable progress has been made toward helping clients develop adequate vocational decision-making skills in terms of three components: seeking information, considering alternatives and consequences, and deciding in terms of success probability. Studies which investigate the principles of reinforcement and social model learning in promoting career exploration and planning have been most prevalent in the literature. Krumboltz and Schroeder (1965) investigated both model reinforcement and reinforcement counseling techniques on the information-seeking behavior of 54 high-school juniors. Their findings indicated that 1) experimental subjects engaged in more information-seeking than controls did; 2) reinforcement counseling was significantly more effective than control procedures with females; and 3) for males, model-reinforcement counseling produced significantly more information-seeking than control procedures.

Ryan (1968) demonstrated that both group reinforcement counseling and evaluated simulation materials (programed vocational materials, films, career-day activities) were significantly more effective than control procedures in promoting vocational decision-making and knowledge of sources of personal data and occupational information. Three hundred junior college subjects, who were below the fiftieth percentile on the SCAT and



ACT and who had unsound or no vocational plans, were enrolled in a program to overcome deficiencies and acquire information about themselves and occupations. Students were randomly assigned to reinforcement group counseling, non-reinforcement group counseling, active control, individual counseling or inactive control groups. The treatments were specified in detail. The major findings were that: 1) subjects in groups receiving reinforcement counseling scored significantly higher on a vocational decision test than subjects receiving group counseling without counseling reinforcement or subjects receiving only individual counseling; 2) subjects in the reinforcement groups scored significantly higher on a test of knowledge of sources of occupational information and personal data than did subjects in the other groups; and 3) subjects in counseling groups using simulation materials and counselor reinforcement scored significantly higher on a vocational decision test than subjects receiving counselor reinforcement without simulation materials. It was interesting that the author controlled for initial decision-making behavior by limiting the population to those whose educational and vocational plans were not definitely decided. The consistency of the findings would indicate that active intervention is an important variable. Since the groups in which counselor cueing and reinforcement techniques were utilized showed greater knowledge of sources of information and performed significantly better on a test of decision-making than did subjects in general group or individual counseling, there are considerable implications for further research. These results raise questions about the sufficiency of a warm, understanding atmosphere for promoting effective decision-making.

Thoresen, Krumboltz and Varenhorst (1967) found that various combinations of models, subjects and counselors, arranged according to sex, influenced the degree of subject career exploration behavior for 96 eleventh-grade students. Four types of social model audio-tape treatments were used: 1) male counselor counseling a male student, 2) male counselor counseling a female student, 3) female counselor counseling a male student, and 4) female counselor counseling a female student. For male subjects, the tape depicting a male counselor and male model was most effective when presented by a male counselor. Sex of the model made little difference for females; the significant variable was the sex of the counselor. Both models were more effective when presented by male. The experimental treatments were significantly more effective in promoting subject career exploration than the inactive control treatment.

Thoresen and Krumboltz (1967) investigated the relationship of specific counselor verbal reinforcements to the client's subsequent career information-seeking behavior. The counselor's verbal reinforcements of the client's report of seeking or his intent to seek information was found to be positively correlated with subsequent client information-seeking outside the

interview. Subjective ratings by clients on the helpfulness of counseling were not significantly associated with the clients' information seeking.

*Additional References:* Fretz and Schmidt (1967); Sullivan (1966); Zunker and Brown (1966).

### *Personal Development and Specific Behavior Change*

Counseling research on personal development and specific behavior change appears to be coming from two camps. One camp is concerned primarily with effecting changes in the self and assumes that the relationship between client and counselor is sufficient to produce these changes, which will subsequently manifest themselves in better generalized functioning outside the counseling situation. The other camp focuses on promoting observable overt behavior changes through specific counseling techniques and assumes that success in using new behaviors as well as feedback from other behaviors will effect changes in the self. The order and direction in which change occurs is a philosophically and empirically open question. Most counselors view the aim of counseling as change (Arbuckle, 1968). Krumboltz (1966) specifies the main concerns toward which counselors promote change; in contrast, Holland (1965, p. 262) simply says that the responsibility of the therapist is to "... recognize, interpret and attempt to change the nature of emotional experience." The review of studies on personal development and specific behavior change indicated that the profession appears to be entering a period of new and creditable research with a consequent closing of the gap between what Patterson (1967) calls relationship and behavior therapy.

### *Self-Exploration*

Researchers have begun to look at how small pieces of the counseling process influence counseling outcomes. The recent work of Truax, Carkhuff and associates deserves special attention in this regard. They have found that high "facilitative conditions" (empathy, positive regard, genuineness, concreteness, and client self-disclosure) were associated with constructive client outcomes and that low levels of these conditions were associated with no change or deteriorative outcomes (Carkhuff, 1967; Carkhuff and Truax, 1966; Truax and Carkhuff, 1963). Holder, Carkhuff, and Berenson (1967) found that counselors were successful in manipulating the level of facilitative conditions offered to clients; the authors also found that high facilitative conditions resulted in significantly greater depth of self-exploration. In the most recent study in this area, Anderson (1968) attempted to look beyond the central core of facilitative conditions to a specific aspect of the therapist's behavior; "confrontation" was defined as the presentation of discrepancies seen by the counselor between the counselor's and the client's views of the same situation. Anderson found that

confrontation was most effective in encouraging client self-exploration when accompanied by high levels of the five facilitative conditions identified above. The consistent finding that high facilitative conditions influence counseling outcomes has clear and important implications for counseling practice, counselor training and research that focuses on specific segments of the counseling process.

Kennedy and Zimmer (1968), attempting to increase the output of self-reference statements, studied the reinforcing value of five commonly used stimuli in a setting simulating an initial counseling interview. Each of three experimenters responded to the statements of 72 female undergraduates in elementary education with "mm-hmm, mm-hmm," an affirming head nod and smile, "good," "I see," or a short paraphrase of the subject's statement. The simulated counseling interview was preceded by a warmup period and followed by an oral examination of awareness. The experimental session consisted of three six-minute periods: free operant baseline, conditioning, and extinction. Experimental subjects were randomly assigned to one of the five stimulus conditions and one of three experimenters. Throughout the session, control group subjects received the stimulus conditions on a random schedule. In the conditioning period for the experimental subjects, the experimenter attempted to respond immediately to every self-reference statement emitted by the subject. In the extinction and free operant periods, the experimenter consistently emitted one of the stimulus conditions at random intervals. The paraphrase and the neutrally toned "mm-hmm" treatments were most effective in promoting self-reference statements; however, the success of these treatments was largely a function of the high rate of operant responding obtained by two of the experimenters.

Ince (1968) also attempted to increase the emission of self-reference statements in a well-controlled laboratory study. Reinforcement was delivered according to fixed-interval schedules (30 seconds, 1 minute, 2 minutes, 3 minutes) after a stable rate of responding was established on a continuous reinforcement schedule. Ince found that the frequency curves for his experimental sample ( $N = 3$ ) clearly indicated that a fixed interval schedule of reinforcement was significantly more effective than no reinforcement or continuous reinforcement schedules. Ince also found that as the subjects' responding increased, the number of counselor reinforcements needed to maintain the behavior decreased. Although these two studies suffer from the artificiality of laboratory experimentation and the possibility that mechanical variables affected the results, they do indicate that the kind and timing of counselor verbal responses influence the kind and degree of self-exploration of the client.

Leib and Snyder (1967) attempted to determine the effect of specific group counseling procedures on underachievement and self-actualization as measured by the Personal Orientation Inventory (Shostrom, 1963). Twenty-



eight underachieving college students who voluntarily dropped out of a psychology reading and study methods course were randomly assigned to either a special lecture session or to one of three group discussion sections. The discussion groups met twice a week, one hour per session, for nine weeks. The discussion leader guided group discussions and interactions on the general topics of motivation, achievement, self-direction, parental communication and resolution of common problems. The 14 subjects in the lecture control met as a group and covered material typically covered in the regular psychology course. The data indicated that there was a significant increase in self-actualization as measured by the self-actualization scale (SA) of the Personal Orientation Inventory within groups as well as significant improvements in GPA. There were no significant differences between treatment groups on either variable.

Branan (1967) studied the effect of the counselor's use of self-experience on client self-disclosure and the increase in self-confidence regarding graduate work. Fifteen male and 15 female graduate students were randomly assigned to three experimental and three active control groups. Each group met weekly for five one-hour sessions. In the three experimental groups the counselor drew attention to and extensively utilized self-experiences. Analyses indicated that the extensive use by the counselor of self-experience statements did not promote significant increases in the subject's amount of self-disclosure or self-confidence or reports of being understood. Post treatment interview data on subjects' perception of the genuineness of the counseling relationship revealed that the control groups scored significantly higher than the self-experience treatment groups. Although the experimental treatment assumed congruence between life styles and experiences of counselor and counsees, the one significant statistical result indicates that disclosure of his self-experience by a counselor may serve as a negative stimulus for promoting genuineness in the counseling relationship. Controlling for sex of counselee and counselor and extending the limits of self-disclosure to include current personal-social experiences of the participants would make an interesting extension of this study.

*Additional References:* Catron (1966); Dicken and Fordham (1967); Kramer (1968); McGreevy and Daane (1967); Neuringer, Myers and Nordmark (1966).

### *Anxiety Reduction*

Counselors and educators have become aware of the debilitating nature of anxiety. For example, highly anxious students appear to receive lower grades and experience a higher rate of academic failure than do non-anxious students of equal intelligence (Spielberger and Weitz, 1964). There is, fortunately, growing evidence that anxiety is amenable to intervention with specific behavior techniques without the occurrence of "relapse" or



"symptom substitution" which traditional "dynamic" theories predict (Paul, 1968).

Emery and Krumboltz (1967) tested the efficacy of group (standard) versus individually tailored anxiety hierarchies in the desensitization of college students concerned with test anxiety. Fifty-four identified test-anxious college freshmen were randomly assigned to either desensitization utilizing an individual anxiety hierarchy, or desensitization with a single standard hierarchy, or a no-treatment control. Data showed that desensitization significantly effected a reduction of test anxiety as measured by a test anxiety scale (a refinement of a test anxiety scale previously constructed by Emery) and self rating by both treatment groups. Final examination grades of experimental and control groups were not significantly different. Further analysis revealed that both desensitization groups experienced anxiety reduction before and during their examinations. No significant differences on any of the criterion measurements were found between subjects utilizing either the standard or individualized hierarchies. The authors concluded that an individualized hierarchy in the desensitization process was not needed under the conditions of the study. The individualized hierarchies, however, were not truly individualized but were a random presentation of the 16 items from the standard hierarchy established in an earlier pilot study by the senior author. Although students in the individualized hierarchy treatment could add items, the initial presentation and task of ranking the standard items precludes considering this as a distinctly separate treatment. As the authors pointed out, the same findings with a third experimental group that constructed their own individualized hierarchies with no initial intervention would have more strongly supported the hypothesis that standard hierarchies are as effective as individual ones in reducing test anxiety.

Clements (1966) evaluated small group counseling in aiding school youth to reduce their anxiety during the transition from high school to college. Clements hypothesized that counseled students would exhibit less anxiety prior and subsequent to college entrance. Sixty students randomly assigned to six subgroups were counseled prior to high school graduation by either of two full-time doctoral students. The two control groups consisted of 120 students who received no group counseling. The sessions focused on the aspects of college attendance that concerned the group members. After the six weeks of group counseling in the spring, individuals in the experimental groups were significantly less anxious about themselves than were the individuals in one control group used for comparison as measured by an index of adjustment and values and a self-concept scale. Differences were maintained subsequent to entrance into college as compared to the second control group.

The contrast between the Emery-Krumboltz and the Clements studies can be seen in the degree of specificity in the counseling process. The

success of both studies in reducing anxiety is possibly explained by Lazarus's (1961) theoretical statement. Lazarus believes that the interview situation is conducive to nonspecific reciprocal inhibition either because verbal interaction can evoke automatic responses similar to deep relaxation or, possibly, because the therapeutic atmosphere of empathy and acceptance can inhibit anxieties.

Paul and Shannon (1966) were successful in producing significant reductions in interpersonal performance anxiety for "chronically" anxious males. Group desensitization and discussion were found to be as effective as individual systematic desensitization and significantly better than insight-oriented psychotherapy and attention-placebo treatments. There was considerable evidence of positive generalization of treatment effects. In addition to using commendable control procedures in this study, Paul (1968) evaluated the long-term effects of the 1966 treatments and found considerable substantiation of the earlier findings. Analysis of self-report measures and the public behavioral criterion of academic success indicated that after two years the significant gains from modified group desensitization were maintained without loss of effectiveness in the treatment of maladaptive anxiety.

*Additional References:* Davison (1968); Kahn and Baker (1968); Katahn, Strenger and Cherry (1966); Thoresen and Neuman (1968).

### *Interpersonal Functioning*

Interpersonal functioning involves the ability to interact successfully with the members of one's environment. Interpersonal performance anxiety, one aspect of this subject, has been discussed in this chapter. However, failure in interpersonal functioning which is caused by inappropriate overt behaviors, poor attending skills, shyness and other correlates of low sociometric status also deserve attention.

Creditable research on the changes in interpersonal functioning which are brought about through counseling has been extended recently to include younger children. Briskin and Gardner (1968) demonstrated that a combination of time-out, nonattention and social reinforcement procedures were effective in reducing inappropriate behaviors and increasing socially acceptable behaviors of nursery school children. The study was completed in a field situation in which no environmental modification was attempted other than the specific strategy applied to the subject under study. On the basis of classroom observations and conferences with the teacher and the subject's mother, socially acceptable or inappropriate behaviors were operationally defined and categorized. In the two-minute time-out procedure, the mother removed her child from the classroom when the teacher signaled that an inappropriate behavior had taken place. Socially acceptable behaviors were strengthened by the teacher's attention and praise. Base-line data indicated that the children spent an average of 31 per cent of

each school session engaged in inappropriate behavior. The treatment procedure resulted in the elimination of much of the previous disruptive behavior. Data taken in the first and second treatment phases and in the follow-up thirty days later showed that the children engaged in socially acceptable behavior from 95 to 98 per cent of the time.

Kennedy and Thompson (1967) used reinforcement techniques and traditional counseling to modify the attending behavior of a first-grade boy during 17 counseling sessions. The time spent on task during the arithmetic period and the completion of lessons were chosen as outcome criteria to determine whether the transfer of training from the conditioning situation to the regular classroom resulted. Treatment produced a 14 per cent increase in time spent on task and nearly 100 per cent increase in lesson completion. In both of the above studies, behavioral improvement cannot be attributed solely to specific treatments, attention must also be given to the increase in the positive reinforcing value of the classroom environment as the child began enjoying some successes in either interpersonal relationships or in academic functioning.

Kranzler et al. (1966) attempted to effect changes in the interpersonal functioning of twenty low-sociometrically-identified fourth graders. The subjects were randomly assigned to either client-centered counseling, teacher guidance or inactive control. In the counseling treatment, eight students met as a group twice a week for six weeks and thereafter in groups of four once a week for 12 weeks. During the second phase each subject also received individual counseling. The four subjects in the teacher-guidance treatment were given positive reinforcement for the work they did, enjoyable classroom tasks to perform and memberships in preferred groups. Neither the controls nor their teachers were aware of their participation in the study. Two sociometric post-tests were administered, the first test at the completion of the five-month study, the second seven months later. The counseled group significantly increased its sociometric status on both tests in comparison to the control group. The teacher guidance group did not differ significantly from either the counseled or control groups.

The Kranzler study is an attempt to bring systematic research to the important area of elementary school counseling. However, a sociometric test might have been administered at the end of the first six weeks of group counseling when the procedure was changed by breaking the counseled group into two groups and adding an individual counseling treatment phase. Moreover, the designers might have anticipated a reverse Hawthorne effect, which may have been the cause of the control group's negative sociometric change. The members of the control group saw that the subjects in both teacher-guidance and counseling treatments were receiving special attention. The significant difference, therefore, may have been produced by a further deterioration of the controls' already unfavorable position rather than to the results of the counseling per se.



In an attempt to correct the methodological weaknesses of the 1966 investigation, Mayer, Kranzler and Matthes (1967) designed a second study in which the subjects wanted to improve their peer relationships. A larger experimental sample and counselors who were trained in elementary counseling were incorporated into the design. Sociometric change scores and teacher ratings of perceptions of their students' popularity before and after the project were utilized as dependent measures. Although the mean sociometric gain score for the subjects in the counseling condition was greater than that for the subjects in either control or teacher-guidance conditions, no significant statistical differences were found. Pre-posttest changes in ranks on the teacher rating procedure also indicated no significant differences between treatments. It is interesting to note, in light of our comments on the previous study, that data on the teacher-guidance treatment indicated negative change scores on both criterion instruments.

*Additional References:* Barclay (1967); O'Leary and Becker (1967).

### Summary

Although it is closing, the persistent gap between research and practice remains wide. Too few studies have been carried out which affect what individuals do as counselors. It may still be hard to refute Steffire's (1963) contention that counselors might learn more about helping adolescents from reading Salinger's *The Catcher in the Rye* than from reading the latest counseling journals.

Most counseling studies still emanate from broad conceptual frameworks. Much of our progress has been attenuated by the failure to design treatments to promote specific outcomes. Any similarity or relationship between counseling treatments and counseling outcomes is too often one of chance. It would seem that researchers should be well beyond doing an experiment in which a treatment, "counseling," is compared to something labeled a control group and in which some outcome criterion, say GPA, is used with no demonstrable relationship to whatever was done in the name of counseling.

Greater consideration of new experimental designs is needed. Counseling treatments designed to promote specific outcomes may need to be multidimensional. Client characteristics, counselor characteristics and combinations of each do affect the outcomes; new designs are needed to test these variables simultaneously with treatment variables. A recent study by Thoresen and Krumboltz (1968), in which the similarity of social models and clients was investigated, exemplifies an attempt to test interaction hypotheses rather than the often used counseling versus no counseling approach. A recent study by Thoresen, Hosford and Krumboltz (1968) is another example in which an interaction design was utilized.



New research designs are needed to avoid the confounding effect of selecting subjects from those who volunteer for counseling. Although researchers who use volunteers only can assign subjects randomly to either experimental or control groups (which is better than comparing volunteers with non-volunteers), the differences between the two groups may be lessened if some control subjects subsequently request counseling. Volunteering and waiting for counseling affects the behavior of many individuals (Volsky et al., 1965) and for some may reduce dissonance (Hill, 1966).

More attention should be given to the goal of preventing problems. The writers of this chapter were unable to find investigations published during 1965-1968 which considered ways or treatments which counselors might use for preventing such things as inappropriate vocational choice, inadequate social relationships and inadequate learning.

The probability that any one study will ever prove counseling effective or ineffective is infinitesimal, and no single study will answer all questions for which the counselor needs answers. A scientific base for counseling will be a greater possibility when researchers use more specific treatments and outcome measures which are appropriate and sensitive enough to detect possible change. If more researchers asked, "Will the results of this study affect what individuals do as counselors?", the gap between research and practice would close appreciably.

### Bibliography

- Anderson, Susan C. "Effects of Confrontation by High- and Low- Functioning Therapists." *Journal of Counseling Psychology* 15: 411-16; 1968.
- Arbuckle, Dugald S. "Counseling Effectiveness and Related Issues." *Journal of Counseling Psychology* 15: 430-35; 1968.
- Bates, Marilyn. "A Test of Group Counseling." *Personnel and Guidance Journal* 46: 749-53; 1968.
- Benson, Ronald L. and Blocher, Don H. Evaluation of Developmental Counseling with Groups of Low Achievers in a High School Setting. *School Counselor* 14: 215-20; 1967.
- Branan, John M. "Client Reaction to Counselor's Use of Self-Experience." *Personnel and Guidance Journal* 45: 568-72; 1967.
- Briskin, Alan S. and Gardner, William I. "Presentation and Removal of Social Reinforcement in Reducing Inappropriate Behavior of a Nursery School Child." *Young Children* 24: 84-89; 1968.
- Bunda, Richard and Mezzano, Joseph. "The Effects of a Work Experience Program on Performance of Potential Dropouts." *The School Counselor* 15: 272-74; 1968.
- Carkhuff, Robert R. *The Counselor's Contribution to Facilitative Processes*. Urbana, Ill.: Parkinson, 1967.
- Carkhuff, Robert R. and Truax, Charles B. "Toward Explaining Success and Failure in Interpersonal Learning Experiences." *Personnel and Guidance Journal* 46: 723-28; 1966.
- Clements, Barton E. "Transitional Adolescents, Anxiety, and Group Counseling." *Personnel and Guidance Journal* 45: 67-71; 1966.
- Emery, John R. and Krumboltz, John D. "Standard Versus Individualized Hierarchies in Desensitization to Reduce Test Anxiety." *Journal of Counseling Psychology* 14: 204-209; 1967.
- Ewing, Thomas N. and Gilbert, William M. "Controlled Study of the Effects of Counseling on the Scholastic Achievements of Students of Superior Ability." *Journal of Counseling Psychology* 14: 235-39; 1967.

- Gilliland, Burl E. "Small Group Counseling with Negro Adolescents in a Public High School." *Journal of Counseling Psychology* 15: 147-52; 1968.
- Ginzberg, Eli. "Toward a Theory of Occupational Choice." *Personnel and Guidance Journal* 30: 491-94; 1952.
- Hill, Arthur H. "Motivation and Academic Counseling." *Journal of Counseling Psychology* 13: 447-53; 1966.
- Holder, Todd; Carkhuff, Robert R.; and Berenson, Bernard G. "Differential Effects of the Manipulation of Therapeutic Conditions Upon High- and Low-Functioning Clients." *Journal of Counseling Psychology* 14: 63-66; 1967.
- Holland, G. A. *Fundamentals of Psychotherapy*. New York: Holt, Rinehart, and Winston, 1965.
- Ince, Laurence. "Effects of Fixed-Interval Reinforcement on the Frequency of a Verbal Response Class in a Quasi-Counseling Situation." *Journal of Counseling Psychology* 15: 140-46; 1968.
- Jackson, Robert M. *Results of Early Identification and Guidance of Underachievers*. Cooperative Research Project No. S-153. LaCrosse: Wisc. State Univ., October 1966.
- Kagan, Norman. "Group Procedures." *Review of Educational Research* 36: 274-87; 1966.
- Kennedy, Daniel A. and Thompson, Ina. "Use of Reinforcement Technique with a First Grade Boy." *Personnel and Guidance Journal* 46: 366-70; 1967.
- Kennedy, John J. and Zimmer, Jules M. "Reinforcing Value of Five Stimulus Conditions in a Quasi-Counseling Situation." *Journal of Counseling Psychology* 15: 357-62; 1968.
- Kiesler, Donald J. "Some Myths of Psychotherapy Research and the Search for a Paradigm." *Psychological Bulletin* 65: 110-36; 1966.
- Kranzler, Gerald D. et al. "Counseling with Elementary School Children: An Experimental Study." *Personnel and Guidance Journal* 44: 944-49; 1966.
- Krumboltz, John D. "Behavioral Goals for Counseling." *Journal of Counseling Psychology* 13: 153-59; 1966.
- Krumboltz, John D. and Schroeder, W. W. "Promoting Career Exploration Through Reinforcement." *Personnel and Guidance Journal* 44: 19-26; 1965.
- Lazarus, Arnold A. "Group Therapy of Phobic Disorders by Systematic Desensitization." *Journal of Abnormal Social Psychology* 63: 504-10; 1961.
- Leib, Jere W. and Snyder, William U. "Effects of Group Discussions on Underachievement and Self-Actualization." *Journal of Counseling Psychology* 14: 282-85; 1967.
- Mayer, G. Roy; Kranzler, Gerald D.; and Matthes, William A. "Elementary School Counseling and Peer Relations." *Personnel and Guidance Journal* 46: 360-65; 1967.
- McDaniel, H. B. *Guidance in the Modern School*. New York: Dryden Press, 1956.
- McGowan, Richard J. "The Effect of Brief Contact Interviews with Low-Ability, Low-Achieving Students." *The School Counselor* 15: 386-89; 1968.
- Patterson, Cecil H. *Relationship Therapy and/or Behavior Therapy*. Paper presented to Annual Convention of the Amer. Personnel Guidance Assoc., March 1967. Urbana: Univ. of Ill., College of Education. (Mimeo.)
- Paul, Gordon L. "Two-Year Follow-Up of Systematic Desensitization in Therapy Groups." *Journal of Abnormal Psychology* 73: 119-30; 1968.
- Paul, Gordon L. and Shannon, Donald T. "Treatment of Anxiety Through Systematic Desensitization in Therapy Groups." *Journal of Abnormal Psychology* 71: 124-35; 1966.
- Roth, Robert M.; Mauksch, Hans O.; and Peiser, Kenneth. "The Non-Achievement Syndrome, Group Therapy, and Achievement Change." *Personnel and Guidance Journal* 46: 393-98; 1967.
- Rothney, John W. M. *Guidance Practice and Results*. New York: Harper and Brothers, 1958.
- Ryan, T. A. *Effect of an Integrated Instructional Counseling Program to Improve Vocational Decision-Making of Community College Youth*. Cooperative Research Project No. HRD 413-65 5-0154. Corvallis: Ore. State Univ., February 1968.
- Ryan, T. A. *Effectiveness of Counseling in College Residence Halls on Students' Study Behavior*. Cooperative Research Project No. 3269. Corvallis: Ore. State Univ., 1967.
- Shostrom, E. L. *The Personal Orientation Inventory*. San Diego: Educational and Industrial Testing Service, 1963.
- Spielberger, Charles D. and Weitz, Henry. *Improving the Academic Performance of Anxious College Freshmen: A Group-Counseling Approach to the Prevention of*

- Underachievement*. Psychological Monographs, General and Applied: Vol. 78, No. 13 (Whole No. 590). Washington, D.C.: Amer. Psychol. Assoc., 1964.
- Steffire, Buford. "Research in Guidance: Horizons for the Future." *Theory into Practice* 2: 44-50; 1963.
- Thelen, Mark H. and Harris, Charles S. "Personality of College Underachievers Who Improve with Group Psychotherapy." *Personnel and Guidance Journal* 46: 561-66; 1968.
- Thoresen, Carl E.; Hosford, Ray E.; and Krumboltz, John D. *Determining Effective Models for Counseling Clients of Varying Competencies*. Stanford: Stanford Univ., 1968.
- Thoresen, Carl E. and Krumboltz, John D. "Relationship of Counselor Reinforcement of Selected Responses to External Behavior." *Journal of Counseling Psychology* 14: 140-44; 1967.
- Thoresen, Carl E. and Krumboltz, John D. "Similarity of Social Models and Clients in Behavioral Counseling: Two Experimental Studies." *Journal of Counseling Psychology* 15: 393-401; 1968.
- Thoresen, Carl E.; Krumboltz, John D.; and Varenhorst, Barbara. "Sex of Counselors and Models: Effect on Client Career Exploration." *Journal of Counseling Psychology* 14: 503-508; 1967.
- Truax, Charles B. and Carkhuff, Robert R. *Recent Advances in the Study of Behavior Change*. Montreal: McGill University Press, 1963.
- Volsky, T. et al. *The Outcomes of Counseling and Psychotherapy*. Minneapolis: Univ. of Minn. Press, 1965.
- Winkler, Ronald C. et al. "The Effects of Selected Counseling and Remedial Techniques on Underachieving Elementary School Students." *Journal of Counseling Psychology* 12: 384-87; 1965.

### Additional References

- Barclay, James R. "Effecting Behavior Change in the Elementary Classroom: An Exploratory Study." *Journal of Counseling Psychology* 14: 240-47; 1967.
- Brown, William F. "Student-to-Student Counseling for Academic Adjustment." *Personnel and Guidance Journal* 43: 811-17; 1965.
- Catron, David W. "Educational-Vocational Group Counseling: The Effects on Perception of Self and Others." *Journal of Counseling Psychology* 13: 202-207; 1966.
- Davison, Gerald C. "Systematic Desensitization as a Counterconditioning Process." *Journal of Abnormal Psychology* 73: 91-99; 1968.
- Dicken, Charles and Fordham, Michael. "Effects of Reinforcement of Self-References in Quasi-Therapeutic Interviews." *Journal of Counseling Psychology* 14: 145-52; 1967.
- Dickenson, Walter A. and Truax, Charles B. "Group Counseling with College Underachievers." *Personnel and Guidance Journal* 45: 243-47; 1966.
- Fretz, Bruce R. and Schmidt, Lyle D. "Comparison of Improvers and Nonimprovers in an Educational Skills Course." *Journal of Counseling Psychology* 14: 175-76; 1967.
- Kahn, Michael and Baker, Bruce. "Desensitization with Minimal Therapist Contact." *Journal of Abnormal Psychology* 73: 198-200; 1968.
- Katahn, Martin; Strenger, Stuart; and Cherry, Nancy. "Group Counseling and Behavior Therapy with Test-Anxious College Students." *Journal of Consulting Psychology* 30: 544-49; 1966.
- Kramer, Howard C. "Effects of Conditioning Several Responses in a Group Setting." *Journal of Counseling Psychology* 15: 58-62; 1968.
- Laxer, R. M. et al. "Counseling Small Groups of Underachievers in Secondary Schools: An Experimental Study." *Ontario Journal of Educational Research* 9: 49-57; 1966.
- LeMay, Morris L. and Christensen, Oscar C. "The Uncontrollable Nature of Control Groups." *Journal of Counseling Psychology* 15: 63-67; 1968.
- McGreevy, C. Patrick and Daane, Calvin J. "Changes in Semantic Differential that Accompany Counseling." *Journal of Counseling Psychology* 14: 526-34; 1967.
- Neuringer, Charles; Myers, Roger A.; and Nordmark, Thorberg, Jr. "The Transfer of a Verbally Conditioned Response Class." *Journal of Counseling Psychology* 13: 208-13; 1966.

- O'Leary, K. Daniel and Becker, Wesley C. "Behavior Modification of an Adjustment Class: A Token Reinforcement Program." *Exceptional Children* 33: 637-42; 1967.
- Stormer, G. Edward. "Milieu Group Counseling in Elementary School Guidance." *Elementary School Counseling and Guidance* 1: 240-54; 1967.
- Sullivan, Howard J. "The Effects of Counseling as a College-Encouragement Procedure." *Journal of Counseling Psychology* 13: 93-97; 1966.
- Thoresen, Carl E. and Neuman, Donald R. *Professional and Subprofessional Counselors Using Group Desensitization and Insight Procedures to Reduce Examination Anxiety*. Stanford: Stanford Univ., 1968. (Mimeo.)
- Zunker, Vernon G. and Brown, William F. "Comparative Effectiveness of Student and Professional Counselors." *Personnel and Guidance Journal* 44: 738-43; 1966.

## AUTHORS

- Hosford, Ray E. Address: University of Wisconsin at Madison, Wisconsin Title: Assistant Professor Age: 35 Degrees: B.A., Central Wash. State; M.A., Central Wash. State; Ph.D., Stanford Univ. Specialization: Behavior Modification Techniques in Counseling; Elementary School Counseling; Social Modeling.
- Briskin, Alan S. Address: University of Wisconsin at Madison, Wisconsin Title: Teaching Assistant; Counselor—B'nai B'rith Vocational Service Age: 35 Degrees: B.S., Marquette Univ.; M.S., Univ. of Wisc. Specialization: Behavior Modification Techniques applied to counseling.





## 6: GROUP COUNSELING

Alan R. Anderson\*

*University of Minnesota*

The rapid growth in the professional practice of group counseling has been accompanied by a proliferation of related research studies. Some 240 of the articles reviewed (covering the period from Summer 1965 to Fall 1968) were considered pertinent to this chapter. Among the studies in Gazda and Larsen's (1968) noncritical abstracts of group counseling research studies in the guidance literature, only those which are particularly relevant to the chapter are repeated here. In this chapter, group counseling was considered to include professional attempts to assist non-psychotic clients to examine and modify their values or behavior through small-group interaction. Although many of the studies were well-designed, there were few long-term treatment studies, follow-up studies, or replication studies. Most of Kagan's (1966) evaluative comments are still valid. The spelling out of specific procedures and techniques in sufficient detail to permit replication, although dimly perceptible, is still essentially lacking. Despite the masses of data being collected and analyzed, most studies are relatively unrelated, small-scale efforts which provide only the accumulation of bits of evidence. There is still no body of theoretically related knowledge on which the practice of group counseling can be solidly grounded. The organization of this chapter represents an

---

\*Dr. George M. Gazda, University of Georgia, and Dr. Merle M. Ohlsen, University of Illinois, served as consultants to Dr. Anderson on the preparation of this chapter. A. Lynn Scoresby, Mary Leonard, Suzanne Bickford and Barbara Anderson also assisted in the preparation of the chapter.

attempt to systematize these bits of evidence. The chapter sections are arranged to follow the order of the concerns a counselor has in setting up and conducting a counseling group: 1) Client Selection and Preparation, 2) Outcomes, 3) Group Interaction, 4) Group Development, 5) Leadership Styles and Group Structure, and 6) Innovations and Suggestions for Research. In the final section, Perspective, an attempt is made to identify trends and summarize some of the chapter highlights.

### *Client Selection and Preparation*

In nearly all studies, the basis for client selection was convenience. In most cases, groups were assembled according to a single common factor such as age or a common problem such as low academic achievement. No studies were found in which the composition of counseling or therapy groups was deliberately and systematically manipulated as a primary variable being investigated—although some studies did produce *post hoc* data.

Mezzano (1967) selected, from each of six counseling groups, two members who had shown the highest degree of "investment" (based on the counselee's willingness to explore his own behavior and communicate openly) and two who had shown the least "investment." Among these 24 clients a positive relationship was found between degree of "investment" and academic improvement. Yalom and Rand (1966) found that highly "compatible" (homogeneous in verbal responses to desired interpersonal areas) psychotherapy groups were significantly more cohesive than less compatible groups. Members who were least "compatible" with their group tended to be less satisfied and to leave their group prematurely. An implication of this study, in light of the suggested relationship between attitudinal similarity and interpersonal attraction (Goldstein, Heller and Sechrest, 1966), is that attitudinally homogeneous groups will generally develop cohesiveness easily and rapidly, but establishing cohesion in heterogeneous groups will require special effort and attention.

Studies of problem-solving groups may be relevant to selecting clients for counseling groups on the basis of personality variables. Goldman, Haberlein and Feder (1965) found that three-member groups of all "conformers" shared ideas and divided work more readily than groups of all "resisters"; the resisters were more hostile and competitive and paid less attention to the ideas of other members. Wallach, Kogan and Burt (1967), studying the "risky-shift" phenomenon, assigned five-member groups of college students to "discuss a topic to consensus." Groups of field-dependent subjects (affect-oriented) reached consensus more quickly and showed a greater increase in their willingness to take risks than did field-independent subjects (analytically and cognitively oriented). This study has relevance for creating a group climate which encourages members to take such risks

as giving negative feedback or disclosing sensitive areas of their own behavior.

Stability of group membership was investigated in two studies. Hall and Williams (1966) compared the decision-making performance of established and *ad hoc* groups under conflict conditions. They found that established groups were superior in performing under conflict conditions. The established groups tended to resolve differences with increased creativity, but the *ad hoc* groups tended merely to work out compromises. Goldberg and Maccoby (1965) investigated the effects of stable versus shifting group membership on the behavior of second-grade children in situations which required that the children interact cooperatively to maximize their goals. Children in the stable groups did significantly better, and there was more of a tendency in the shifting groups for high-scoring members to exercise dominance over low-scoring members.

Controversy over optimal group size rages endlessly, but few attempts are made to study the effects of group size on specific process or outcome variables. From their review of research on group size relative to problem-solving groups, Goldstein, Heller and Sechrest (1966) observed that as group size increases: 1) the absolute rate of interaction for any given member tends to decrease; 2) the proportion of infrequent contributors to the group interaction increases; 3) more members report feelings of threat and inhibition regarding participation; 4) giving of information and suggestions increases and asking for opinions and showing agreement decreases; and 5) more statements are directed to the leader and the leader addresses more statements to the group as a whole rather than to individual members. Goldstein, Heller and Sechrest suggested ending impressionistic speculation and conducting studies of group size related to variables such as nature and rate of verbal interaction and leader interventions.

There were no reports found of studies in which an attempt was made to prepare clients systematically for optimal participation in their group. The procedures designed to prepare a client for his group experience varied from simply sending an invitation to a student to join an academic improvement group to the elaborate steps followed by Bach (1954) which involved extensive testing, several interviews, reading, and trial attendance at a group session. Data obtained from studies of group cohesiveness (Cartwright and Zander, 1960) suggest the potential productivity of research on the effects of establishing the client's expectations that group experiences will be attractive and valuable or that certain kinds of interaction will be likely to develop among group members. Goldstein, Heller and Sechrest (1966) suggested conducting studies which stem from the hypothesis that people tend to "find" more value in experiences in which they have invested considerable energy. Further hypotheses might be formulated relating motivation and expectancies to interaction norms and outcomes; these hypotheses then could form the basis for developing specialized procedures



designed to prepare specific clients for participation in specific groups.

Group counseling research reflects little interest in client selection or client preparation as major independent variables. The available data suggest that people who are affectively oriented, flexible, highly motivated to change, and sufficiently well adjusted to interact rationally with others function well in counseling groups. Often, however, if anticipated results are not achieved, it is prematurely concluded that group counseling is not helpful for "this kind of client." Client selection, complex as it is, must be studied in relation to both process and outcome variables. Specifically, composing groups on the basis of predicted compatibility relative to preferred style of interaction appears most promising. Research is needed which will help to predict how a given client will respond under a given set of conditions and with a particular combination of other group members, including the leader.

*Additional Reference:* Thelen and Harris (1968).

### *Outcomes*

Although it may be said that all research focuses ultimately on outcomes, this section includes only those studies concerned with achieving specified outcomes through group counseling. Campbell and Dunnette (1968) have suggested three "disarmingly simple" standards for evaluating T-group experiences which, with minor modifications, could be applied to outcome research on counseling groups. The standards are: 1) measures of clients which broadly sample goal-related behavior should be explicitly determined before and after the counseling experience; 2) measured client changes during the time that counseling occurred should be compared with similar, uncounseled persons; and 3) effects of completing the evaluation instruments should be studied. A fourth standard is that the nature of the treatment, the specific procedures used, should be spelled out explicitly in behavioral terms.

### *Academic Performance*

Attempts to improve the grade point averages (GPA) of students were reported in most of the outcome studies. Few studies were found which approached the standards proposed by Campbell and Dunnette. Roth, Mauksch and Peiser (1967) showed that subjects in mandatory counseling groups of 7 to 12 members raised their mean GPA significantly from .94 to 1.74, compared to control subjects whose GPA did not change. The control group, however, was not well matched with the experimentals on either academic aptitude or prior academic achievement. Despite an attempt to account for the regression effect by comparing changes in the GPA of the high- and low-achieving students in both groups, replication with a tightened research design is necessary before definitive conclusions

can be drawn. Rationale and treatment procedures are given in sufficient detail that replication is possible. Mezzano (1968) studied the immediate and ten-week effects of group counseling and group-plus-individual counseling on randomly assigned low-achieving high-school students. No differences were found on GPA, study attitudes or self-concept immediately following counseling—but, ten weeks later the counseled groups had significantly higher mean GPAs than control groups had.

Dickenson and Truax (1966) suggested that frequent equivocal results of group counseling with underachievers may be a result of lumping together high and low "therapeutic conditions." They matched groups of neurotic, underachieving college freshmen; the freshmen were then counseled under high and moderate therapeutic conditions (accurate empathy, nonpossessive warmth and genuineness offered by the group leader). Significant improvement in mean GPA was found for the subjects who received the high therapeutic conditions, but not found for those in the moderate or control groups.

Although GPA is a socially significant, objectively quantifiable, convenient measure of academic success, it is somewhat inadequate as a single outcome criterion. It is generally acknowledged that many kinds of academic learnings are not reflected in the GPA; this suggests that GPA should only be used as one of several criteria of academic success. In addition, lumping together the various behaviors which contribute to GPA precludes the investigation of the specific behaviors which may be affected by the treatment.

### *Attitude and Personality Change*

Three studies in this section approached the standards proposed by Campbell and Dunnette. These three studies used comparable procedures (desensitization) which were operationally defined, were concerned with one independent variable (anxiety), used multiple criteria and had adequate control groups.

Paul and Shannon (1966) provided systematic desensitization of anxiety for two five-member groups of highly anxious students who had been members of a delayed treatment control group in a previous study. The treatment consisted of training the subjects in deep muscle relaxation; constructing a hierarchy of anxiety producing situations; and beginning with the least anxiety-producing situations in the hierarchy, counter-conditioning the anxiety by directing the subjects to "image" the events while in a state of deep muscle relaxation. The images proceeded from least to most threatening items, determined by specific rules of timing. The results indicated that anxiety (as measured by a test battery consisting of anxiety inventories, observer ratings and self-reports) decreased significantly after the group desensitization treatment; the anxiety scores had not changed during the previous semester when no treatment was provided. Subjects

in the desensitization groups also had a significantly higher mean GPA at the end of the semester than did the matched control subjects.

Katahn, Strenger and Cherry (1966) used two groups of six and eight students each to assess the effect of systematic group desensitization combined with group discussion. Anxiety scores, as measured by a modification of Sarason's Test Anxiety Scale, decreased significantly in the two desensitization groups after six weeks; there was no significant decrease of anxiety scores in the control groups, which were composed of similar students whose schedules were such that they were not included in the treatment program. The mean GPA of the therapy groups also increased significantly compared to the control group. The study by Paul and Shannon and the study by Katahn, Strenger and Cherry provide rather conclusive evidence that the treatment used was effective in reducing anxiety and improving the academic performance of the highly anxious college students. However, a survey given to the subjects in the latter study indicated that the clients were unanimous in the opinion that it was the informal discussion (also used in the Paul and Shannon study), rather than the desensitization procedures, which was responsible for the results.

Thoresen and Neuman (1968) compared systematic group desensitization (without informal discussion) with group insight counseling in reducing examination anxiety among 54 college undergraduates. Groups of three were led by professional (Ph.D.'s with clinical experience) or subprofessional (first year, inexperienced graduate students in psychology) counselors for five weekly sessions. The group desensitization treatment was found to reduce anxiety, as measured by three self-report measures of anxiety, significantly more than group insight treatment or two control group procedures. Differences were not found between these treatments on an observer checklist; both the group desensitization treatment and the group insight treatment were significantly more effective than control procedures. The subprofessional counselors who administered the treatments were, in general, as effective as the professional counselors. These findings support the efficacy of systematic group desensitization without discussion. However, the investigators found that some subjects in the desensitization groups reduced their anxiety more than others; this was also true for some subjects in the insight groups. Studies are now needed to determine what client characteristics interact with which treatments (or combination of treatments) to produce the most effective, efficient change.

Attempts to evaluate the outcomes of laboratory groups such as T-groups have been conducted in recent years. Space limitations preclude their inclusion here. Campbell and Dunnette (1968) have provided an excellent critical review of research in that field.

## *Career Development*

Despite the appearance of a substantial body of theoretical and research literature on career development (Tennyson, 1968), there have been few recent attempts to facilitate vocational planning through group counseling. Gilliland (1968) randomly assigned Negro high-school students either to all-male and all-female groups for "group-centered" counseling or to a control group. After one year the counseled students showed significant gains in reading ability, English usage, occupational aspiration and vocational development; the control subjects exhibited no significant gains. Treatment procedures were not specified. Jesse and Heimann (1965) tested the effects of group guidance on the vocational maturity of ninth-grade boys. Data from a vocational maturity scale and from interviews indicated no significant change in vocational maturity by the following summer. Specification of treatment procedures, lacking in both of these studies, is necessary for valid conclusions to be drawn.

Stewart and Thoresen (1968) reported two studies with eighth- and eleventh-grade students. Some students in the sessions listened in groups of four to three audio-presented group social models; each presentation was followed by group reinforcement discussion. Other students engaged in reinforcement discussion after completing written materials on topics such as "Possible Interests After High School," "Getting Helpful Information," and "Occupation and Me." Although subjects in the experimental treatments were significantly more engaged, on the average, in information-seeking behaviors than control subjects were, considerable individual variations were found. These findings suggested the need for multivariate studies in which interactions between treatments and subject characteristics are explicitly examined.

Thoresen, Hosford and Krumboltz (1968) used a multivariate design in four high schools to study interactive effects of social model and subject characteristics on career relevant information-seeking behaviors of 189 male high-school juniors. In one group session, one of nine different audio social models, representing three levels of either academic, social or athletic success, was played to matched groups of seven Ss. Each group was matched on level (high, moderate, low) of academic, athletic or social success. Results were mixed in that the high-success athletic model, for example, was significantly more effective with highly and moderately successful athletic subjects in one school, but was ineffective in a second school. Significant interactions were also found among counselors, schools, subjects and treatments; this finding suggested that some counselors using particular model treatments with certain subjects were more effective than other counselors. This study merits replication, but more attention should be given to how group counselor characteristics and increased treatment sessions affect outcomes.

A number of viable career development models have appeared in



recent years; they are amenable to implementation through group counseling (Super et al., 1963; Tiedeman and O'Hara, 1963). Specific group procedures attempting to engender behaviors relevant to those theories might be tested on selected groups of clients. More attempts to induce vocational maturity through some generalized form of group counseling are clearly not needed. Vocational maturity and group counseling are not single, clearly understood entities. What is needed are multivariate projects which utilize multiple outcome criteria and differential treatments.

*Additional References:* Anderson and Johnson (1968); Pattison (1965); Gundlach (1967); Shapiro and Birk (1967).

### *Group Interaction*

Of all the factors which impinge upon a counseling group, none is more pervasive than the interaction among group members. Ultimately, the interactive-communication in a group determines the outcomes. Clear specification of treatment procedures requires systematic identification of the information contained in a unit of communication, the manner in which the information is given, and the effect it has on the relationship between the receiver and sender of the message (Watzlawick, Beavin and Jackson, 1967). The purpose of group interaction research is to classify systematically and evaluate the communication occurring in a group.

Truax (1968) and Truax and Carkhuff (1967) presented evidence from a number of studies which indicated that changes in group participants usually occurred when a high degree of genuineness, warmth and empathic understanding was being communicated within the group. The evidence also suggested that such changes in group participants were unlikely to occur in the absence of the communication of such relationship qualities. The authors found that the leader's use of these qualities (genuineness, warmth, etc.) to reinforce selectively the self-exploratory statements in counseling groups increased the frequency of subsequent self-exploratory statements. These qualities also influenced outcome criteria such as improved personal and interpersonal adjustment (measured by personality inventories) and the length of time that patients spent outside the mental hospital during the subsequent year. The therapeutic gains in the above studies were attributed primarily to relationship qualities communicated by the group leader.

Working with groups of high-school students, Hansen, Zimpfer and Easterling (1967) found significant correlations between increased congruence of self-concept and the perceptions which group members had of the relationships offered by other group members. The authors found no correlations between congruence of self-concept and the members' perceptions of relationships offered by the counselor. Unfortunately, these

studies have not provided sufficient detail about treatment procedures to allow replication. In addition, it is not possible to attribute specific outcomes to specific kinds of interaction because of the gross description of process and outcomes.

A number of attempts to assess systematically or to classify group interaction have been reported. Mann (1967) developed an interaction scoring system based on psychoanalytic theory and derived from observation of a series of relatively unstructured social psychology classes. Eight of the 16 categories describe members' affective (impulse) responses to the group leader; three categories describe feelings activated by the leader's status in the perceived authority structure; and five categories describe how a member felt about himself in relation to the leader. There were also four levels of relating to the leader, dependent upon the directness of the communication.

The Hill Interaction Matrix (Hill, 1965), developed from observations of psychotherapy groups in a mental hospital, is a two-dimensional system based on Bion's (1959) classification of verbal interaction. The four horizontal categories were derived from the content of the communication; they are *references to topics of general interest, the group, problems and concerns of individual members, and relationships of group members to each other in the "here and now."* The five vertical categories are derived from the "work-style," which might also be called the process or relationship dimension. The three "pre-work" categories are *responsive, conventional and assertive*; the two "work" categories are *speculative and confrontive*. Superimposing the five work-styles over the four content styles creates a twenty-cell matrix into which any statement can be classified according to content and work level.

In a recent study of five-member T-groups composed of college students who volunteered for the study, Dunnette and Campbell (1969) found that after six two-hour sessions there was a significant positive correlation between the level of interpersonal communication, as measured by the Hill Interaction Matrix, and the increase in empathic understanding.

There are a number of constructs in the group dynamics field which, if operationalized, could be fruitfully studied in terms of group interaction. For example, Truax (1961) reported significant relationships between group cohesiveness and degree of patient self-exploration and patient insight. Goldstein, Heller and Sechrest (1966) presented evidence from non-therapy groups which showed that group cohesion is a powerful influence on individual member behavior; they proposed a series of testable hypotheses to relate cohesiveness to a number of specific, therapy-related variables. As molar concepts such as norms, cohesiveness, trust, empathy, warmth and congruence are operationalized and quantified within interaction systems, it becomes possible to relate them to counseling outcomes for

individual clients in predictable ways. Few studies approach this degree of rigorous specificity.

*Additional References:* Bonney (1965); Hoffman, Burke and Maier (1965); Ohlsen and Pearson (1965).

### *Group Development*

Eight years ago, Hill (1961) listed some 34 "theories" of group development drawn from group dynamics and group psychotherapy literature. Nearly all of the theories were formulated from intuitive observations and couched in the language of each author's theoretical bias. Tuckman (1965), in criticizing group development studies for lack of experimental rigor, suggested a classification scheme based on setting, realm (group structure and task activity) and developmental sequence. Tuckman hoped that this classification scheme would stimulate additional research.

Mann (1967), studying the development of four social psychology classes on group process, suggested a "cyclical" rather than a "phase" model of group development. These groups, rather than moving "ever onward and upward" to maturity, experienced repeated ups and downs. Mann further observed that a shift in the predominant tone of the group was rarely shared by all members equally. The emergence of subgroups appeared to be primarily responsible for the group's cycles, and an attempt was made to identify the issues which created the conditions under which the subgroups emerged, competed and combined, thereby altering the direction of the group's development. Fiebert (1968) observed that leader behavior in a sensitivity group changed from that of a *catalyst* (one who encourages and facilitates) to that of an *orchestrator* (one who directs verbal and non-verbal exercises and deepens explorations of feeling and motives) and then to that of a *participant* (one who takes a member role and blends into the group). Findings by Lubin and Zuckerman (1967) lend some support to Fiebert's observations; they found that the level of anxiety, depression and hostility in a one-week sensitivity training conference rose sharply about halfway through (when the orchestrator was deepening the exploration of feelings) and then declined.

Scott (1965) studied two groups of delinquents assigned to group therapy by the juvenile court. Scott used seven observable characteristics of interaction to identify four discrete phases of group development. Sufficient detail was provided for replication, using differently composed groups or groups meeting under different circumstances. Ahearn (1968) used the Hill Interaction Matrix to study the interaction patterns in a group of ten prospective counselors who held 56 sessions. Ahearn found that the number of statements considered to be of high therapeutic value tended to increase sharply from the first to second quarter; then the number of statements

leveled off. Intermediate level interaction decreased in the second quarter and then increased in the last two quarters. This "polarization" is described as a work-flight pattern in which the group alternates between work and tension-release.

Group development is the one area of small-group work in which theoretical models abound but a solid empirical base is lacking. Mann's study illustrates the difficulty and expense of thoroughly studying group development. Each group meeting over a period of time is an  $n$  of one. It is likely that the patterns of interaction which develop in a counseling group will depend upon the group composition, leadership style and expected outcomes; however, the precise nature of these variations in relation to those factors is yet to be studied. Such studies are generally feasible only in large institutional settings or through the collaboration of several investigators.

*Additional References:* Heckel et al. (1967); Foley and Bonney (1966).

### *Leadership Styles and Group Structure*

This section contains a review of studies in which two or more "leadership styles" are compared or in which a particular mode of leadership behavior is used as an independent variable. The differences in styles ranged from gross variations such as group-centered versus leader-centered, to very specific variations such as conscious verbal reinforcements of different particular types of interaction.

Thoresen and Krumboltz (1967) examined the relationships between five categories of statements recorded in individual and group sessions and the amount of information-seeking behavior outside the group. They found significant positive correlations of 1) the frequency of counselor verbal reinforcement of client information-seeking responses, 2) the frequency of client information-seeking responses and 3) the number of counselor "cue" responses, with outside information-seeking behaviors. Subjective client rating of the degree of "helpfulness" of the counseling were unrelated to the frequency of information-seeking behavior engaged in by clients.

Truax (1968) demonstrated that mental hospital patients in groups receiving high levels of leader behaviors (described as accurate empathy, nonpossessive warmth, and genuineness) showed significant increases in frequency of self-exploratory responses; the same patients also improved personal adjustment and spent more time out of the hospital during the subsequent year. Truax concluded, as a result of this and related studies, that the "therapeutic-triad" of accurate empathy, nonpossessive warmth, and genuineness could appropriately be construed as reinforcing patient self-exploration and extratherapy improvement.



Culbert (1968) studied the effects of leader self-disclosure in two T-groups of ten college students each; the groups met twice weekly in two-hour sessions for fourteen weeks. One of the weekly sessions was spent in the T-group with two co-trainers; the other session was spent in a dyad pairing with another group member. Culbert found that although an equivalent number of two-person "perceived therapeutic relationships" were formed in each group, subjects in the group with less-disclosing leaders more often entered into relationships with leaders and dyad partners, whereas the subjects with more self-disclosing leaders entered more often into relationships with other group members. Although both groups reached the same level of self-awareness, the group in which the leaders were more self-disclosing reached their peak level sooner.

Salzberg (1967), in a study of the effects of the presence or absence of a leader, examined an open-ended group composed of 35 inpatients of a privileged psychiatric ward; the inpatients participated at various times during a 16 week period. No significant difference in total frequency of statements made was found between sessions attended by the therapist and those sessions which the therapist observed but did not attend. More comments were classified as "spontaneous" when the therapist was absent, but fewer comments were classified as "relevant"; this result suggests that the group members worked harder on individual problems when the leader was present. Seligman's (1968) study produced similar results. Leaderless groups in Seligman's study spent significantly more time on "low-level" interaction (measured by the Hill Interaction Matrix) than leader-led groups spent; the reverse finding was true for "high-level" interaction.

The amount of structure optimally provided by the leader is a subject of considerable controversy in group counseling—despite evidence from group dynamics which suggests that the need for structure in a group depends upon the situation. Gilbreath (1967) found that leader-structured groups were more effective in increasing the ego strength and GPA of freshmen and sophomore male underachievers than were non-directive or group-structured groups. Gilbreath also found that the subjects with high dependency needs (measured by a second order factor of the Sterns Activity Index) tended to raise their GPA when they participated in leader-structured groups; dependent subjects participating in non-directive groups did not tend to raise their GPA. In contrast, independent subjects improved their GPA in the non-directive groups, but showed no improvement in GPA in the leader-structured groups.

The weakness of most of these studies lies in the researchers' vagueness about the specific procedures involved in the various treatments. The strength of these studies lies in the researchers' attempt to test the differential effects of a given process variable on different kinds of clients.

*Additional References:* Bates (1968); Delaney and Heimann (1966); French, Sherwood and Bradford (1966); Shaw and Blum (1966).

## *Innovations and Trends*

Presenting a section on innovations involves the risk of being perceived either as "behind the times" or as "faddish" and unscientific. Nevertheless, exciting things are happening in group work, and many of them will have great impact on future practice and research. Human-encounter groups of various kinds and labels are becoming popular in churches, neighborhoods, schools, hospitals and businesses. "Self-help" groups, encouraged by the success of Alcoholics Anonymous, were given a kind of scientific respectability by Mowrer's *The New Group Therapy* (1964) and by the Integrity Therapy movement. Groups such as Alanon for families of alcoholics and Alateens for children of alcoholics are becoming common throughout the country. There are also many half-way houses for alcoholics. Synanon, a half-way house organization for narcotics users; Recovery, Inc., for former mental patients and excessively nervous people; and TOPS, for weight-watchers, are examples of self-help groups which use a variety of group procedures.

Esalen (Murphy, 1967) represents the focal point of a relatively new movement attempting to "expand the mind of man" through encounter groups and non-verbal individual and group experiences. Esalen attracts both the "hippies" and the elite; it has been extolled by some and condemned by others. Nonetheless, Esalen appears destined to have a substantial impact on counseling and therapy—particularly in group settings. Sense Relaxation (Gunther, 1968) and Joy (Schutz, 1967) deal with concepts and experiences designed to stimulate man to a greater sense of his own existence in relation to that of others. These works presage what appears to be the major thrust in Humanistic Psychology.

One of the more conventional innovations in group counseling and therapy is the *marathon*, in which groups meet either continuously or with short breaks for 12 to 72 hours. These groups may be composed of people looking for relief from conflict and loneliness or of people who are essentially happy but are seeking new ways to enrich their interpersonal encounters and their own self-awareness. Bach (1967, p. 995) has described marathon groups as the "most direct, the most efficient, and the most economical antidote to alienation, meaninglessness, fragmentation and other hazards of mental health in our time."

Gazda (1968) has compiled a number of accounts of innovative approaches to group therapy; he has included marathons, video-tape feedback, the TORI process and others. Otto (1967) has presented a theory and program for developing human potential through specified group activities; Boocock and Schild (1968) have collected a series of simulated games for use in a variety of groups. Bessell (1968) is organizing a Human Development Program in which personal confidence and desired classroom behaviors are learned through modeling and rehearsal of these behaviors in brief, small-group sessions held daily in the classroom.

The evidence regarding the precise nature and impact of these innovations is either lacking or is based on intuitive observation. Hopefully, each innovation will be investigated with sufficient rigor that it can be systematically defined and then either refined or discarded. These innovations represent one of the frontiers of knowledge in group counseling. In one sense, they differ from many of the "tried and true" ways only because they are new, but not because there is any less evidence regarding their efficacy. In another sense, they differ radically from standard approaches. These innovations are not bound by the standard constructs which place heavy emphasis on verbal communication. Man is viewed more comprehensively, as a biological-psycho-social system, and non-verbal activities such as music, dance, color and games are utilized to induce imagery, fantasy, bodily contact, introspection and experimentation—which appear to influence behavior more dramatically than talk strategies.

The changes that result from such non-verbal activities (dance, music, etc.) are often referred to as experiential learnings. Can such learnings be isolated and related to specific kinds of stimuli? Do they last, or do they simply serve as temporary facilitators of more permanent learnings? Do these learnings tend to generalize and transfer to normal life situations or are they existential? Do these learnings interfere with defense mechanisms in a psychologically dangerous manner, or is it possible to predict the impact of a particular experience on a given client and thus to construct experiential programs which will teach the desired behaviors? Does the very act of asking such questions inhibit the existential impact of the experience? Hopefully, such questions will be answered through rigorous inquiry and experimentation. It is relatively unexplored territory for the group researcher.

*Additional Reference:* Lazarus and Bienlein (1967).

### *A Perspective*

Counseling is increasingly viewed as an influence process, and counselors are exploring more active ways of helpfully intervening in the lives of their clients. There are indications that the field of group counseling is maturing. The amount of research activity is increasing rapidly, and there is liberal dissemination of information through professional journals. The series of desensitization studies is encouraging because it represents a successful effort to replicate procedures and tighten research design. In the desensitization studies, the treatment procedures and the nature of the clientele and the outcomes were precisely specified. This specificity is also evident in the operant conditioning studies. Solid evidence is accumulating that operant conditioning; desensitization procedures; and the presence of accurate empathy nonpossessive warmth and genuineness in a

relationship tend to produce consistent, predictable outcomes with selected clients.

Despite Gundlach's (1967, p. 205) conclusion that "there is no simple, universal patient; there is no universal treatment named group therapy; and there is no simple, wonderful, universal outcome measure," many studies still seem to be searching for "the truth." Much of what is published is inconsequential because it contributes little or nothing to the pool of knowledge about group counseling. In most studies, researchers used gross, non-operational descriptions of treatment procedures (or none at all) and single criteria such as a GPA. There was a common failure in most studies to specify why the treatment that was used should be expected to produce the predicted outcomes with the particular clientele studied. Efforts were seldom made to study the differential effects of treatment procedures on different clients.

Group counseling needs studies that specify concrete, measurable goals for individual clients, detailed analysis of verbal and non-verbal communication which constituted the treatment, and a variety of appropriate criteria. The formulation of general principles about group counseling still awaits the use of identical treatment procedures in different settings with different clientele, as well as multivariate approaches which compare several treatment procedures in similar settings with similar clientele (Cohn, 1967). Immediate and long-term treatment effects must also be studied.

Theoretical models, broad enough to span both counseling theory and group dynamics (and perhaps the humanities), need to be formulated; broadly conceived research methodology must be developed to test these models. Gazda's (1968) compilation of basic approaches to group counseling may provide a beginning. However, it is doubtful that sufficient experimental rigor can always be provided in natural settings. Experiments on "quasi-counseling" groups are needed as interim steps to provide necessary controls and to reduce the risk to clients. Institutional or collaborative research that is based on these theoretical models and within which small coordinated studies can be conducted appears to hold the most promise for the future.

### Bibliography

- Ahearn, Thomas R. *An Interaction Process Analysis of Extended Group Counseling with Prospective Counselors*. Doctor's thesis. Athens: Univ. of Ga., 1968.
- Bach, George R. *Intensive Group Psychotherapy*. New York: Ronald Press, 1954.
- Bach, George R. "Marathon Group Dynamics." *Psychological Reports* 20: 995-999; 1967.
- Bessel, Harold. "The Content is the Medium: The Confidence is the Message." *Psychology Today* 1: 32-35, 61; January 1968.
- Bion, W. R. *Experiences in Groups*. New York: Basic Books, 1959.
- Boocock, Sarane S. and Schild, E. O. *Simulation Games in Learning*. Beverly Hills: Sage, 1968.



- Campbell, John and Dunnette, Marvin D. "Effectiveness of T Group Experiences in Managerial Training and Development." *Psychological Bulletin* 79: 73-104; 1968.
- Cartwright, Dorwin and Zander, Alvin. *Group Dynamics: Research the Theory*. Evanston, Ill.: Row, Peterson, 1960.
- Culbert, S. A. "Trainer Self-Disclosure and Member Growth in Two T Groups." *Journal of Applied Behavioral Science* 4: 47-73; 1968.
- Cohn, Benjamin (editor). *Guidelines for Future Research on Group Counseling in the Public School Setting*. Washington, D. C.: Amer. Personnel and Guidance Assoc., 1967.
- Dickenson, Walter A. and Truax, Charles B. "Group Counseling with Underachievers." *Personnel and Guidance Journal* 45: 243-247; 1966.
- Dunnette, Marvin D. and Campbell, John. "People Feeling: Joy, More Joy and the Slough of Despond." *Journal of Applied Behavioral Science* 5, No. 1, in press.
- Fiebert, Martin S. "Sensitivity Training: An Analysis of Trainer Intervention and Group Process." *Psychological Reports* 22: 829-838; 1968.
- Gazda, George M. (editor). *Basic Approaches to Group Psychotherapy and Group Counseling*. Springfield, Ill.: C. C. Thomas, 1968.
- Gazda, George M. (editor). *Innovations to Group Psychotherapy*. Springfield, Ill.: C. C. Thomas, 1968.
- Gazda, George M. and Larsen, Mary Juhan. "A Comprehensive Appraisal of Group and Multiple Counseling Research." *Journal of Research and Development in Education* 1: 57-132; 1968.
- Gilbreath Stuart H. "Group Counseling, Dependence, and College Male Underachievement." *Journal of Counseling Psychology* 14: 449-453; 1967.
- Gilliland, Burl E. "Small Group Counseling with Negro Adolescents in a Public High School." *Journal of Counseling Psychology* 15: 147-153; 1968.
- Goldberg, Moses H. and Maccoby, Eleanor E. "Children's Acquisition of Skill in Performing a Group Task under Two Conditions of Group Formation." *Journal of Personality and Social Psychology* 2: 898-902; 1965.
- Goldman, Morton; Haberlein, Bernard J.; and Feder, Gloria J. "Conformity and Resistance to Group Pressure." *Sociometry* 28: 220-226; 1965.
- Goldstein, Arnold P.; Heller, Kenneth; and Sechrest, Lee B. *Psychotherapy and the Psychology of Behavior Change*. New York: John Wiley and Sons, 1966.
- Gundlach, Ralph H. "Outcome Studies in Group Psychotherapy." *International Journal of Group Psychotherapy* 17: 196-210; 1967.
- Gunther, Bernard. *Sense Relaxation: Below Your Mind*. New York: Macmillan, 1968.
- Hall, Jay S. and Williams, Martha S. "A Comparison of Decision Making Performance in Established and Ad Hoc Groups." *Journal of Personality and Social Psychology* 3: 214-222; 1966.
- Hansen, J. C.; Zimpfer, D. G.; and Easterling, R. E. "A Study of the Relationships in Multiple Counseling." *Journal of Educational Research* 60: 461-463; 1967.
- Hill, William F. "Group Development." *Group Psychotherapy: Collected Papers in Practice, Theory and Research*. Provo: Utah State Hospital Press, 1961.
- Hill, William F. *Hill Interaction Matrix*. Los Angeles: Univ. of So. Calif. Press, 1965.
- Jesse, Bill A. and Heimann, Robert A. "Effects of Individual and Group Counseling on Vocational Maturity of Ninth Grade Boys." *Journal of Educational Research* 59: 68-73; 1965.
- Kagan, Norman. "Group Procedures." *Review of Educational Research* 36: 274-287; 1966.
- Katahn, Martin; Strenger, Stuart; and Cherry, Nancy. "Group Counseling and Behavior Therapy with Test Anxious College Students." *Journal of Consulting Psychology* 30: 544-49; 1966.
- Lubin, Bernard and Zuckerman, Marvin. "Affective and Perceptual-Cognitive Patterns in Sensitivity Training Groups." *Psychological Reports* 21: 363-376; 1967.
- Mann, Richard Dewey. *Interpersonal Styles and Group Development: An Analysis of the Member-Leader Relationship*. New York: Wiley, 1967.
- Mezzano, J. "A Consideration for Group Counselors: Degree of Counselee Investment." *The School Counselor* 14: 167-169; 1967.
- Mezzano, J. "Group Counseling with Low-Motivated Male High School Students—Comparative Effects of Two Uses of Counselor Time." *Journal of Educational Research* 61: 223-224; 1968.
- Mowrer, O. Hobart. *The New Group Therapy*. Princeton, N. J.: D. Van Nostrand, 1964.

- Murphy, Michael. "Esalen: Where It's At." *Psychology Today* 1: 34-39; December 1967.
- Otto, Herbert A. *Guide to Developing Your Potential*. New York: Charles Scribner's Sons, 1967.
- Paul, Gordon L. and Shannon, Donald T. "The Treatment of Anxiety Through Systematic Desensitization in Therapy Groups." *Journal of Abnormal Psychology* 71: 124-135; 1966.
- Roth, Robert M.; Mauksch, Hans O.; and Peiser, Kenneth. "The Non-Achievement Syndrome, Group Therapy, and Achievement Change." *Personnel and Guidance Journal* 46: 393-398; 1967.
- Salzberg, Herman C. "Verbal Behavior in Group Psychotherapy With and Without a Therapist." *Journal of Counseling Psychology* 14: 24-27; 1967.
- Schutz, William C. Joy. New York: Grove Press, 1967.
- Scott, Max. *Group Development: An Exploratory Study of Small Group Growth Patterns*. Master's Thesis. Provo, Utah: Brigham Young Univ., 1965.
- Seligman, Milton. *Verbal Behavior in Therapist-Led, Leaderless and Alternating Group Psychotherapy Sessions*. Doctor's thesis. Eugene: Univ. of Ore., 1968.
- Stewart, Norman R. and Thoresen, Carl E. *Behavioral Group Counseling for Career Development*. Symposium presented at Amer. Personnel and Guidance Assoc. Convention in Detroit, 1968.
- Super, Donald E. et al. *Career Development: Self Concept Theory*. New York: College Entrance Examination Board, 1963.
- Tennyson, Wesley W. "Career Development." *Review of Educational Research* 38: 346-366; 1968.
- Thoresen, Carl E. and Krumboltz, John D. "Relationship of Counselor Reinforcement of Selected Responses to External Behavior." *Journal of Counseling Psychology* 14: 140-144; 1967.
- Thoresen, Carl E. and Neuman, Donald R. *Professional and Subprofessional Counselors Using Group Desensitization and Insight Procedures to Reduce Examination Anxiety*. Stanford, Calif.: School of Education, Stanford Univ., 1968. (Mimeo.)
- Thoresen, Carl E.; Hosford, Roy E.; and Krumboltz, John D. "Determining Effective Models for Counseling Clients of Varying Competencies." Stanford, Calif.: School of Education, Stanford Univ., 1968. (Mimeo.)
- Tiedeman, David V. and O'Hara, Robert P. *Career Development: Choice and Adjustment*. New York: College Entrance Examination Board, 1963.
- Truax, Charles B. *The Process of Group Psychotherapy: Relationships between Hypothesized Therapeutic Conditions and Intrapersonal Exploration*. Psychological Monographs: General and Applied: Vol. 75, No. 7 (Whole No. 505). Washington, D. C.: Amer. Psychol. Assoc., 1961.
- Truax, Charles B. "Therapist Interpersonal Reinforcement of Client Self-Exploration and Therapeutic Outcome in Group Psychotherapy." *Journal of Counseling Psychology* 15: 225-231; 1968.
- Truax, Charles B. and Carkhuff, Robert R. *Toward Effective Counseling and Psychotherapy: Training and Practice*. Chicago: Aldine, 1967.
- Tuckman, Bruce W. "Developmental Sequence in Small Groups." *Psychological Bulletin* 63: 384-399; 1965.
- Wallach, Michael A.; Kogan, Nathan; and Burt, Roger B. "Group Risk Taking and Field Dependence-Independence of Group Members." *Sociometry* 30: 323-338; 1967.
- Watzlawick, Paul; Beavin, Janet H.; and Jackson, Don D. *Pragmatics of Human Communication*. New York: W. W. Norton, 1967.
- Yalom, Irvin D. and Rand, Kenneth. "Compatibility and Cohesiveness in Therapy Groups." *Archives of General Psychiatry* 15: 267-275; 1966.

### Additional References

- Anderson, Alan R. and Johnson, Donald L. "Using Group Procedures to Improve Human Relations in the School System." *School Counselor* 15: 334-342; 1968.
- Bates, Marilyn. "A Test of Group Counseling." *Personnel and Guidance Journal* 46: 749-753; 1968.
- Bonney, Warren C. "Pressures Toward Conformity in Group Counseling." *Personnel and Guidance Journal* 43: 970-973; 1965.

- Delaney, Daniel J. and Heimann, Robert A. "Effectiveness of Sensitivity Training on the Perception of Non-Verbal Communications." *Journal of Counseling Psychology* 13: 436-440; 1966.
- Foley, W. J. and Bonney, W. C. "A Developmental Model for Counseling Groups." *Personnel and Guidance Journal* 44: 576-590; 1966.
- French, John R. P., Jr.; Sherwood, John J.; and Bradford, David L. "Change in Self-Identity in a Management Training Conference." *Journal of Applied Behavioral Science* 2: 210-218; 1966.
- Gundlach, Ralph H. "Outcome Studies in Group Psychotherapy." *International Journal of Group Psychotherapy* 17: 196-210; 1967.
- Heckel, Robert V. et al. "Emergence of Distinct Verbal Phases in Group Therapy." *Psychological Reports* 21: 630-632; 1967.
- Hoffman, L. Richard; Burke, Ronald J.; and Maier, Norman R. F. "Participation, Influence, and Satisfaction Among Members of Problem-Solving Groups." *Psychological Reports* 16: 661-67; 1965.
- Lazarus, Herbert R. and Bienlein, Donna K. "Soap Opera Therapy." *International Journal of Group Psychotherapy* 17: 252-256; 1967.
- Ohlsen, Merle M. and Pearson, Richard D. "A Method for the Classification of Group Interaction and Its Use to Explore the Influence of Individual and Role Factors in Group Counseling." *Journal of Clinical Psychology* 21: 436-441; 1965.
- Pattison, E. M. "Evaluation Studies of Group Psychotherapy." *International Journal of Group Psychotherapy* 15: 382-397; 1965.
- Shapiro, D. and Birk, L. "Group Therapy and Experimental Perspective." *International Journal of Group Psychotherapy* 17: 211-224; 1967.
- Shaw, Marvin E. and Blum, J. Michael. "Effects of Leadership Style Upon Group Performance as a Function of Task Structure." *Journal of Personality and Social Psychology* 3: 238-242; 1966.
- Thelen, M. H. and Harris, C. J. "Personality of College Underachievers Who Improve with Group Psychotherapy." *Personnel and Guidance Journal* 46: 561-566; 1968.

#### AUTHOR

Anderson, Alan R. Address: University of Minnesota at Minneapolis, Minnesota  
 Title: Assistant Professor Age: 38 Degrees: B.S., Univ. of Alberta; B.Ed., Univ. of Alberta; M.S., Brigham Young Univ.; Ph.D., Brigham Young Univ. Specialization: Group Counseling; Human Relations Laboratory Training.

## 7: CAREER DEVELOPMENT

**John L. Holland\***

*American College Testing Program*

**Douglas R. Whitney**

*American College Testing Program*

During the last three years an increasing number of people, professions and organizations have concerned themselves with vocational training, access to vocational training and jobs, vocational counseling, vocational choice, and careers. This increased interest and its associated financial support have attracted so many new investigators to the field of vocational behavior and stimulated the work of others to the extent that a comprehensive review is no longer possible. Consequently, only the more limited area of career development is treated in this review.

The research published during the period of April 1965 to July 1968 was reviewed for this chapter. However, the chapter authors have included only those publications which they considered to be useful scientific contributions and genuine elements of career development (vocational theory, vocational choice, work history). The articles and papers which were selected include speculation about these behaviors and the determinants, predictions, classifications, and patterns of vocational preferences for various intervals of time. For earlier, closely related reports, the reader should see reviews by Perrone (1966) and Tennyson (1968).

---

\*Dr. John O. Crites, University of Iowa; Dr. Anne Roe, Harvard University; and Dr. Buford Steffire, Michigan State University, served as consultants to Drs. Holland and Whitney on the preparation of this chapter.



## Theory and Speculation

The continuing proliferation of attempts to explain vocational behavior have inspired several authors to collect, organize and evaluate the most influential theories of career development. The most comprehensive work was done by Crites (1968), whose voluminous text *Vocational Psychology* deals with vocational choice, adjustment, success, theory, satisfaction, etc. In *Theories of Career Development*, Osipow (1968) reviewed and evaluated five popular theoretical orientations (Roe, Holland, Ginsberg, Psychoanalytic and Super); Osipow also offered his own orientation and noted some deficiencies in present theories. Carkhuff, Alexik, and Anderson (1967) used a diagrammatic-philosophical standard similar to theories of natural science to evaluate the same theoretical orientations (plus Tiedeman's) that Osipow reviewed. The brief evaluation by Carkhuff, Alexik and Anderson was a mixture of a few perceptive observations, some theoretical misunderstandings and much selective reporting. The works of Crites and Osipow are particularly valuable because the authors organized, interpreted, and evaluated a great range of writing and empirical results in single, extensive sources. All three publications, however, offer more suggestions for researchers than for counselors. In addition, the divergence of opinion illustrates the role that personal orientation plays in both the construction and evaluation of theories. For example, although these authors employed similar evaluative criteria, they did not agree on which were the best or worst theories, on whether or not a particular theory was worthy of consideration, or on a common set of recommendations. Generally these authors neglected the need to tolerate some ambiguity in the process of evolving theories and the equally important need to encourage diverse frameworks for understanding vocational behavior.

Individual investigators have continued to restate and elaborate their ideas. In a recent paper, Super (1968) provided a lucid account of the evolution and current status of his thinking about vocational development. Super continued to stress the need to incorporate broad sources of knowledge into the construction of theories of vocational development. He labeled his approach "differential-developmental-social-phenomenological psychology." The longitudinal study which served as a major vehicle for the testing and revision of Super's theory is reported under "longitudinal studies." In a related study, Crites (1965) developed a useful measure of vocational maturity called the Vocational Development Inventory (VDI).

In a two-volume work, Tiedeman and Dudley (1967) discussed a wide range of topics and problems, including a review of their recent work in career development at Harvard, a statement of conceptual orientation, a description of an information system for vocational decision and a discussion of the professional implications of their work. Specialists and scholars will find these reports the best single source for learning about the development and current status of Tiedeman's orientation. However, students will

often find the diverse technical and esoteric topics in Tiedeman's work beyond their comprehension.

In 1966, Holland (1966) revised his theory of vocational choice, referring to it as "a theory of personality types and model environments." Holland's second statement was more systematic and inclusive because he added explicit definitions for the major concepts and extended the theory to cope with personal development and other behavior, but the theory still contains ambiguities, an incomplete classification scheme and many complex and unnecessary concepts.

Several investigators have focused on the decision-making process for understanding vocational behavior. Katz (1966) made perhaps the most detailed attempt to apply a decision model to the guidance process. He proposed to bring "the student's values explicitly into the decision-making process" and to combine the student's values with the available alternatives and the probabilities of success and reward. A substantial project incorporating Katz's ideas into a vocational guidance system for two-year colleges is now under way. In the same area, Hershenson and Roth (1966) proposed a two-dimensional model of vocational development focused on narrowing the range of possibilities and strengthening the remaining possibilities.

Ellis and Tyler (1967) developed and examined the usefulness of "a system for classifying the . . . factors persons consider in making occupational decisions" (p. 6). They proposed their scheme as an accounting and analytical device to serve two purposes: 1) to discern "the varied modes of orientation that youth from different socioeconomic backgrounds have to . . . work," and 2) to account systematically for "the 'accidental' factor in occupational choice" (p. 39). The value of this work for future research applications was not clear.

In a sociological analysis, Kuvlesky (1966) outlined a longitudinal model of the occupational placement process and its interaction with the development of occupational orientations from childhood to old age. This viewpoint, similar in many ways to Super's orientation, is a useful antidote for the intuitive use of terminology and the parochialism of psychologists.

*Additional References:* Berdie (1966); Chansky (1967).

### *Vocational Images, Preferences, and Influences*

The literature concerning people's perceptions of vocations, the personal occupational choices and the kinds of vocational interests that are associated with certain kinds of occupational preferences continued to grow, but at a decreasing rate.

#### *Vocational Images*

Student images of occupations were found by O'Dowd and Beardslee

(1967) to remain relatively consistent throughout the college years. Furthermore, the images seemed to be common among students in different programs and grade levels. This major report, their second, is a valuable supplement to an earlier cross-sectional study of occupational images (O'Dowd and Beardslee, 1960). In another study, Guilford (1967, p. 64) conducted an inverse factor analysis, the results of which clearly implied that "divergent occupational groups (a) can be differentiated empirically, (b) fit the stereotypes held by the group members, and (c) conform exceedingly well to stereotypes hypothesized on the basis of previous research findings."

### *Vocational Influences and Preferences*

The correlates and presumed determinants of occupational choices remained a popular topic during the last three years. Large sample studies by Werts (1966, 1968) explored some of the relationships among student ability, social class, and parental occupation. Other investigations by Baruch, Segal, and Handrick (1967) and by Johnson (1967) explored the possible influence of family and religion in occupational preference. These studies provide valuable knowledge, but they need to be integrated with theories of vocational development.

In a cross-sectional study, James (1965) explored the way in which commitment to an occupation influenced a person's self-expectations and concept of his chosen occupation. James's study, undertaken to test a nonvocational theory of change, was one of a smaller number of isolated but promising efforts which should be assimilated into more traditional vocational theory.

Routine studies of the differences between students or adults preferring or working in various occupations continued unabated. The studies by Campbell and Schuell (1967), Anderson and Barry (1965), and Ivey and Peterson (1965) illustrate some of the most useful of these studies. Again, an effort to integrate these studies into some theoretical scheme might make them more useful.

### *Occupational Classification*

The classification of occupations remained an unappreciated problem, despite its potential value for understanding vocational behavior. In a longitudinal study of 804 men, Roe et al. (1966), employing Roe's eight-category classification scheme, found that when men changed jobs 68 per cent moved to jobs in the same major group. This and other findings strongly implied that a person's work history is more lawful than many believe it to be. Closely related articles by Hutchinson and Roe (1968) and by Osipow (1966) reinforce the usefulness of the Roe classification system

for arranging occupations and occupational preferences within and between occupational groups.

Subsequently, Holland and Whitney (1968), using Holland's revised classification, found that a longitudinal study of the occupational preferences of male college students conformed to theoretical expectations with a high level of predictive efficiency. For example, Holland and Whitney (1968, p. 16) stated that "79% of the men indicate successive vocational choices that are related or lawful rather than random." A special classification scheme for women yielded similar and equally efficient predictions.

In an interesting departure from traditional classification studies, Goldschmid (1967) scaled 55 academic fields on two dimensions—science and humanities—and then determined by multiple regression how personality traits correlated with choice of major. Although this experiment was moderately successful, its practical and theoretical potentials cannot be determined until more work is done.

Taken together, these studies on occupational classification promise much. The efficiency of the predictions obtained by Roe et al. (1966) and Holland and Whitney (1968) equal or exceed those obtained by more complex and expensive procedures. Equally important, the theoretical rationales for both schemes provide relatively direct and simple explanations of the vocational behavior under consideration.

### *Longitudinal Studies*

Recent years have been characterized by a large number of diverse, constructive, well-planned longitudinal studies of vocational preferences, interests and careers. The substantive outcomes of this work are usually clear, but the theoretical contributions are sometimes obscure because the relationships between the empirical analyses and prior theoretical statements are often ambiguous.

### *Theoretical Studies*

Perhaps the most comprehensive and substantive study among these reviewed was the *Influence of Identity Processes on Student Behavior and Occupational Choice* (Krullee, O'Keefe and Goldberg, 1966). This four-year study of college students at Northwestern provided information about the influences that lead to occupational choice and changes. In this study, the relationship between theories of identity and experimental activities was tenuous; that is, ideas from identity theory guided the design and data collection, but only a few explicit hypotheses about student identity were tested.

Super, Kowalski and Gotkin (1967) reported a ten-year study of Super's theorizing about vocational development. This elaborate study used special criteria developed to assess Super's theoretical concepts. Like many



theoretical studies, this one was marred by unfriendly data—the investigators found that widely used standard measures in educational and vocational guidance, such as intelligence or parental occupation, were better predictors of vocational development in young adulthood than the theoretical measures developed especially for that purpose. Other investigators, Gribbons and Lohnes (1966) and Gribbons, Halperin and Lohnes (1966) examined Super's ideas in longitudinal studies and applied, with some success, a stochastic model (the Markov chain) to the vocational development of high school and college students.

Holland (1968) tested the usefulness of his revised theory (Holland, 1966) in a large, diverse sample of college students. Generally, the majority of hypotheses about personality types received statistically significant support. The hypotheses about student-college interactions received only weak and sometimes ambiguous support. For example, congruent student-college interactions were rarely found to be conducive to the students' satisfaction with college and stability of vocational choice.

Astin and Panos (1968) used several large scale national samples to study the influence of college characteristics upon a student's choice of occupation and major field. Because the study by Astin and Panos included both personal and environmental data as well as some estimates of the relative contribution of each to vocational decisions, the results of the study were especially helpful in analyzing the influence of college environments as they relate to vocational decision-making.

Cooley (1967) used Project TALENT data to examine the interactions among interests, abilities, and career plans. Subsequently, Cooley and Lohnes (1968) developed the career development tree, a method for showing the patterns of successive occupational preferences or occupations, and applied this model to the five-year data in Project TALENT. Their model appeared to be potentially valuable, but critics surely will observe that they cannot see the data because of all those trees.

Using the voluminous longitudinal data in the Strong archives, Campbell et al. (1967) developed a set of basic interest scales (scored according to content such as adventure, agriculture, art, etc.) for the Strong Vocational Interest Blank. The application of these new scales to many specific occupations over long periods of time provided useful information about the relation of vocational interests to occupational status.

### Surveys

The work reported here is a mixed bag of useful survey results and focused and diffuse research analyses. Although these surveys were often atheoretical, their informational value frequently exceeded that of more pretentious studies.

In his book *Undergraduate Career Decisions*, Davis (1965) methodically organized and interpreted data from a large national sample of

college students. In a related article comparing the level of scholastic competition at different colleges with student career decisions, Davis (1966) found that a student judges himself more by comparisons with other students at his college than in some absolute way. For example, a student's decision to give up pre-law may depend in part on what college he attends.

More recently, Sharp, Jones, and Krasnegor (1967) surveyed a national sample of college students five years after graduation. This extensive survey clearly indicated the marked stability of occupational preference during this five-year period. In another survey, Miller (1968) used census data to examine the relationship between formal training and current occupation. Doyle (1965) studied the careers of alumni from a single college. These three studies revealed that careers after college graduation are relatively stable and predictable from plans at graduation.

Several surveys and investigations explored the transition from high school to work. In studies by Heath and Strowig (1967) and Droege (1968), the authors attempted with some success to validate many predictors, including the GATB, against educational and occupational criteria. Partially reported studies of the transition from high school to work by Hoyt (1968) and by Michie and Soldahl (1968) provide the kind of information about students and jobs that many investigators have discussed but seldom gathered: how students find jobs and what they think about their training. These studies are particularly important because they provide previously missing information about the student-work transition with unusual practicality and completeness.

*Additional References:* Little (1967); Kuvlesky and Bealer (1967).

### *One Evaluation*

The years 1965-1968 were marked by a rapid growth of research activity concerned with career development, but there has been no information explosion if only the findings and ideas worth knowing are counted. Although progress is being made, the need for substantive knowledge and better theoretical pictures looms large; to use an agricultural analogy, the soil is well-worked but only a few crops are growing well or are near harvest.

A larger and better integration of current knowledge is needed. The potential value of looking at vocational behavior in the context of counseling, industrial psychology, and particularly sociology needs more explicit attention. Despite much talk about careers, only a few investigations have actually been concerned with careers or work histories. Most investigators have been content to study the relationship between one job and the next or the correlates of a single job. Counseling psychologists have focused on adolescents; industrial psychologists have focused on selection and placement; and sociologists have focused on work groups and other group

phenomena. Consequently, the links among these diverse fields—all concerned with vocational behavior—are few and weak. More long-term longitudinal studies in which concepts from many fields are integrated in single studies are perhaps the most promising possibility.

Fortunately, there are some beginnings. Both Katzell (1964) and Vroom (1964) showed, in theoretical terms, that vocational choice is analogous to job satisfaction, job participation, and turnover. In a series of integrated, theoretical reports concerned primarily with work adjustment, Dawis, Lofquist, and Weiss (1968) expressed similar ideas. The latter reports are valuable for two reasons: they explicitly define concepts and interrelationships, and their operational definitions are practical and ready for new investigation. Relying on his diverse research on vocational behavior, Ginzberg (1966) recently constructed a new developmental theory for understanding patterns of career development and value orientations.

In *Organizational Careers*, Glaser (1968) provided a source of valuable ideas about careers because the readings he has included deal with a host of socially relevant ideas rarely studied by psychologists. Glaser's book is a comprehensive, well-organized group of reports by some of the best American sociologists. Unfortunately, Glaser's cohorts ignore work done by psychologists, a reflection of the mutual lack of exchange between psychologists and sociologists which inhibits the development of a comprehensive understanding of vocational behavior. In short, some tentative, plausible ideas exist that students and researchers might exploit as stepping stones to better understandings.

In addition to the need for more diverse reading and speculation, some other reorientations are desirable. The first reorientation might be called the need for open-mindedness in research and training; the second, the need to reduce destructive researchmanship.

The need for open-mindedness shows itself in many ways, but most explicitly when researchers talk about theory. They usually act as if the philosophy of science and associated research strategies were a closed, explicit body of knowledge complete with how-to-do-it kits, rather than an open, ambiguous collection of ideas and methods. Because researchers have relied on dogmatic views of the philosophy of science, a constricted view of science usually limits their scientific activities. The current climate in education and psychology of a narrow experimentalism, indiscriminate statistical manipulations and timid theorizing further reduces both the range of speculation and the variety of approaches to new knowledge. To illustrate this point, who at this time would be caught performing a simple atheoretical, useful predictive study just because it seemed like a good idea? Psychologists have forgotten their scientific heritage. Much of what is known has come from a great range of activities: practical, theoretical, accidental, experimental, and even occult. Some antidotes which a few readers may find constructive and reassuring include Kaplan's (1964) *The Conduct of Inquiry* and Koestler's (1964) *The Act of Creation*.



The current orientation has been effectively passed to students and instead of a problem-oriented, open-minded breed of researchers, there are too many investigators with a narrow view of what is permissible scientific activity. In his address, "Fads, Fashions, and Folderol in Psychology," Dunnette (1966) provided a witty account of self-destructive researchmanship. The *Sociological Imagination* by Mills (1959) is a larger work devoted to similar problems of research and the researcher. The reader of this literature may find the solace and the necessary perspective to carry on, but it is unlikely that the problems of researchers will change much as long as we live in the American culture with its emphasis on entrepreneurship!

### Bibliography

- Anderson, Harry E. and Barry, John R. "Occupational Choices in Selected Health Professions." *Personnel and Guidance Journal* 44: 177-184; 1965.
- Astin, Alexander W. and Panos, Robert J. *The Educational and Vocational Development of American College Students*. Washington, D. C.: American Council on Education, 1968.
- Baruch, Rhoda; Segal, Stanley; and Handrick, Fannie. *Career and Family: Variations on a Theme*. Harvard Studies in Career Development, No. 52. Cambridge, Mass.: Center for Research in Careers, Harvard Univ., January 1967. (Mimeo.)
- Campbell, David P. and Schuell, Hildred. "The Vocational Interests of Women in Speech Pathology and Audiology." *Asha* 9: 67-72; March 1967.
- Campbell, David P. et al. *A Set of Basic Interest Scales for the Strong Vocational Interest Blank for Men*. Minneapolis: Center for Interest Measurement Research, Univ. of Minn., 1967.
- Carkhuff, Robert R.; Alexik, Mae; and Anderson, Susan. "Do We Have a Theory of Vocational Choice?" *Personnel and Guidance Journal* 46: 335-345; 1967.
- Chansky, Norman M. *A Preliminary Study of the Career Systems Inventory*. Philadelphia: Temple Univ., 1967. (Mimeo.)
- Cooley, William W. *Interactions Among Interests, Abilities, and Career Plans*. Journal of Applied Psychology Monographs: Vol. 51, No. 5 (Whole No. 640). Washington, D. C.: Amer. Psychol. Assoc., 1967.
- Cooley, William W. and Lohnes, Paul R. *Predicting Development of Young Adults*. Cooperative Research Project No. 3051, Interim Report No. 5. Palo Alto, Calif.: Project TALENT Office, Amer. Institutes for Research and Univ. of Pittsburgh, 1968.
- Crites, John O. *Measurement of Vocational Maturity in Adolescence: I. Attitude Test of the Vocational Development Inventory*. Psychological Monographs: General and Applied, Vol. 79, No. 2 (Whole No. 595). Washington, D. C.: Amer. Psychol. Assoc., 1965.
- Crites, John O. *Vocational Psychology*. New York: McGraw-Hill, 1968.
- Davis, James A. *Undergraduate Career Decisions*. Chicago: Aldine, 1965.
- Davis, James A. "The Campus as a Frog Pond: An Application of the Theory of Relative Deprivation to Career Decisions of College Men." *American Journal of Sociology* 72: 17-31; July 1966.
- Davis, René V.; Lofquist, Lloyd H.; and Weiss, David J. *A Theory of Work Adjustment (A Revision)*. Minnesota Studies in Vocational Rehabilitation No. 23. Minneapolis: Industrial Relations Center, Univ. of Minn., 1968.
- Doyle, Robert E. "Career Patterns of Male College Graduates." *Personnel and Guidance Journal* 44: 410-415; 1965.
- Droege, Robert C. "GATB Longitudinal Validation Study." *Journal of Counseling Psychology* 15: 41-47; 1968.
- Dunnette, Marvin D. "Fads, Fashions, and Folderol in Psychology." *American Psychologist* 21: 343-352; 1966.



- Ellis, Robert A. and Tyler, Leona. *Planned and Unplanned Aspects of Occupational Choices by Youth: Toward a Morphology of Occupational Choice*. Eugene: Univ. of Ore., 1967. (Mimeo.)
- Ginzberg, Eli. *The Development of Human Resources*. New York: McGraw-Hill, 1966.
- Glaser, Barney G. (editor). *Organizational Careers*. Chicago: Aldine, 1968.
- Goldschmid, Marcel L. "Prediction of College Majors by Personality Tests." *Journal of Counseling Psychology* 14: 302-308; 1967.
- Gribbons, Warren D. and Lohnes, Paul R. *Career Development*. Cooperative Research Project No. 5-0088. Weston, Mass.: Regis College, 1966.
- Gribbons, Warren D.; Halperin, S.; and Lohnes, Paul R. "Applications of Stochastic Models in Research on Career Development." *Journal of Counseling Psychology* 13: 403-408; 1966.
- Gullford, Joan S. "Isolation and Description of Occupational Stereotypes." *Occupational Psychology* 41: 57-64; 1967.
- Heath, B. R. G. and Strowig, R. Wray. "Predicting Occupational Status for Non-College-Bound Males." *Personnel and Guidance Journal* 46: 144-149; 1967.
- Hershenson, David B. and Roth, Robert M. "A Decisional Process Model of Vocational Development." *Journal of Counseling Psychology* 13: 368-370; 1966.
- Holland, John L. *The Psychology of Vocational Choice*. Waltham, Mass.: Blaisdell, 1966.
- Holland, John L. *Explorations of a Theory of Vocational Choice: VI. A Longitudinal Study Using a Sample of Typical College Students*. *Journal of Applied Psychology Monographs*, Vol. 52, No. 1, Part 2. Washington, D. C.: Amer. Psychol. Assoc., 1968.
- Holland, John L. and Whitney, Douglas R. *Changes in the Vocational Plans of College Students: Orderly or Random?* ACT Research Reports, No. 25. Iowa City: American College Testing Program, Research and Development Division, April 1968.
- Hoyt, Kenneth B. *Transition from School to Work*. Paper prepared for the Executive Director, National Manpower Policy Task Force. Iowa City: Div. of Counselor Education, Univ. of Iowa, 1968. (Mimeo.)
- Hutchinson, Thomas and Roe, Anne. "Studies of Occupational History: Part II. Attractiveness of Occupational Groups of the Roe System." *Journal of Counseling Psychology* 15: 107-110; 1968.
- Ivey, Allen E. and Peterson, Mark B. "Vocational Preference Patterns of Communications Graduates." *Educational and Psychological Measurement* 25: 849-856; 1965.
- James, Fleming, III. "Occupational Choice and Attitude Change." *Journal of Counseling Psychology* 12: 311-315; 1965.
- Johnson, Benton (editor). *Planned and Unplanned Aspects of Occupational Choices by Youth: Religion and Occupational Preference*. Eugene: Univ. of Ore., 1967. (Mimeo.)
- Kaplan, Abraham. *The Conduct of Inquiry: Methodology for Behavioral Sciences*. San Francisco: Chandler, 1964.
- Katz, Martin. "A Model of Guidance for Career Decision-Making." *Vocational Guidance Quarterly* 15: 2-10; 1966.
- Katzell, Raymond A. "Personal Values, Job Satisfaction, and Job Behavior." *Man in a World at Work*. (Edited by Henry Borow.) Boston: Houghton Mifflin, 1964. Chapter 15, pp. 341-363.
- Koestler, Arthur. *The Act of Creation*. New York: Macmillan, 1964.
- Krulse, Gilbert K.; O'Keefe, Robert; and Goldberg, Martin. *Influence of Identity Processes on Student Behavior and Occupational Choice*. Cooperative Research Project No. 5-0809. Evanston, Ill.: Northwestern Univ., 1966.
- Kuvlesky, William P. *The Social-Psychological Dimensions of Occupational Mobility*. Paper presented to the National Vocational Technical Education Seminar, April 1966. Raleigh: Center for Research, Training, and Occupational Education, North Carolina State Univ. (Mimeo.)
- Michie, Jack and Soldahl, Thomas A. *Dominant Factors Influencing the Employment Success of a Selected Group of Disadvantaged Youth. Follow-up Study Class of 1965*. Oakland Unified School District Report No. 3. Oakland, Calif.: the District, Department of Counseling and Occupational Information, 1968. (Mimeo.)
- Miller, Ann R. *Current Occupation and Past Training of Adult Workers*. Office of Statistical Standards, Statistical Evaluation Report No. 7, Bureau of the Budget. Washington, D. C.: Superintendent of Documents, Government Printing Office, 1968. (Mimeo.)

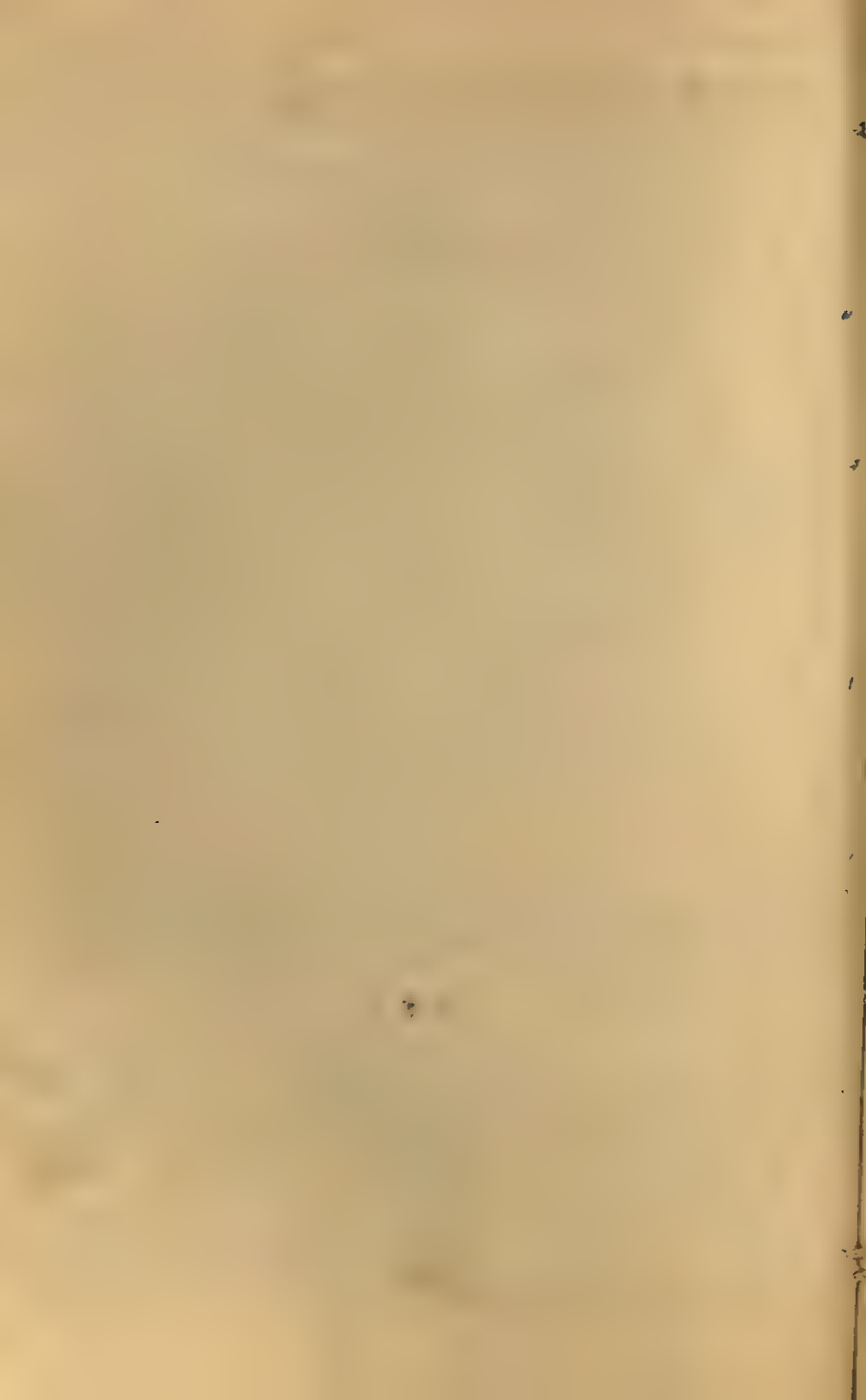
- Mills, C. Wright. *The Sociological Imagination*. London: Oxford Univ. Press, 1959.
- O'Dowd, Donald D. and Beardslee, David C. *College Student Images of a Selected Group of Professions and Occupations*. Cooperative Research Project No. 562(8142). Middletown, Conn.: Wesleyan Univ., 1960.
- O'Dowd, Donald D. and Beardslee, David C. *Development and Consistency of Student Images of Occupations*. Cooperative Research Project No. 5-0858. Rochester, Mich.: Oakland Univ., 1967.
- Ospow, Samuel H. "Consistency of Occupational Choices and Roe's Classification of Occupations." *Vocational Guidance Quarterly* 14: 285-286; 1966.
- Ospow, Samuel H. *Theories of Career Development*. New York: Appleton-Century-Crofts, 1968.
- Perrone, Philip A. "Vocational Development." *Review of Educational Research* 36: 298-307; 1966.
- Roe, Anne et al. "Studies of Occupational History: Part I. Job Changes and the Classifications of Occupations." *Journal of Counseling Psychology* 13: 387-393; 1966.
- Sharp, Laure M.; Jones, Richard C.; and Krasnegor, Rebecca B. *Five Years After the College Degree*, Vol. 1-5. Washington, D. C.: Bureau of Social Science Research, 1965-67.
- Super, Donald E. *Vocational Development Theory: Persons, Positions, and Processes*. Paper presented to the Washington Univ. Symposium on Vocational Development Theory, July 1968. New York: Teachers College, Columbia Univ. (Mimeo.)
- Super, Donald E.; Kowalski, Ruth S.; and Gotkin, Elizabeth H. *Floundering and Trial after High School*. Career Pattern Study, Monograph IV. Cooperative Research Project No. 1393. New York: Teachers College, Columbia Univ., 1967.
- Tiedeman, David V. and Dudley, Gordon A. *Thought, Choice and Action: Processes of Exploration and Commitment in Career Development*, Vol. 1-2. Cambridge, Mass.: Information System for Vocational Decisions, Harvard Graduate School of Education, 1967. (Offset.)
- Tennyson, W. Wesley. "Career Development." *Review of Educational Research* 38: 346-366; 1968.
- Vroom, Victor H. *Work and Motivation*. New York: John Wiley and Sons, 1964.
- Werts, Charles E. *Career Choice Patterns: Ability and Social Class*. NMSC Research Reports, Vol. 2, No. 3. Evanston, Ill.: National Merit Scholarship Corporation, 1966.
- Werts, Charles E. "Paternal Influence on Career Choice." *Journal of Counseling Psychology* 15: 48-52; 1968.

## Additional References

- Berdie, Ralph F. *Vocational Interests and Personality*. Minneapolis: Office of the Dean of Students, Univ. of Minn., 1966. (Mimeo.)
- Kuvlesky, William P. and Bealer, Robert C. "The Relevance of Adolescents' Occupational Aspirations for Subsequent Job Attainments." *Rural Sociology* 32: 290-301; 1967.
- Little, J. Kenneth. "The Occupations of Non-College Youth." *American Educational Research Journal* 4: 147-153; 1967.

## AUTHORS

- Holland, John L. Address: Box 168, Iowa City, Iowa Title: Vice-President, Research and Development Age: 49 Degrees: B.A., Univ. of Omaha; Ph.D., Univ. of Minnesota Specialization: Vocational Behavior; Higher Education.
- Whitney, Douglas R. Address: American College Testing Program, Iowa City, Iowa Title: Research Assistant Age: 25 Degrees: B.S., Iowa State Univ.; M.A., Ohio Univ.



## 8: COUNSELING STUDENTS WITH SPECIAL PROBLEMS

D. David Island\*

*University of Washington*

In this chapter, research reported in the past three years on counseling students with special problems is reviewed. Problems termed social-emotional and educational have served as the primary focus, although some attention was given to physical and intellectual problems. A "special problem" was defined as a category which described certain behavior observed in a relatively small proportion of students. "Special problem" was not used to indicate a rigid, unique classification of individuals. In this way, the interaction between an individual's behavior and his environment was emphasized instead of stressing an assumed underlying trait or state.

Thus, research on counseling college-bound students, counseling women, or vocational counseling, for example, was not searched and reviewed, but research on vocational counseling with deaf students or counseling students with attendance problems was considered. Experimental research articles and studies pertaining to counseling students with special problems made up the bulk of the material reviewed. Descriptive studies, theoretical articles, or descriptions of guidance and counseling programs are not reviewed here.

---

\*Dr. Kenneth B. Hoyt, University of Iowa, served as consultant to Dr. Island on the preparation of this chapter.



## *Special Social and Emotional Problems*

Research on counseling students who have special social or emotional problems was unbalanced and scattered. Research was either scant or nonexistent on counseling students with many kinds of special social or emotional problems. For example, counseling for racial and cultural disadvantage or difference and the effects of poverty or affluence received little attention. Student nonconformity, apathy, alienation, or militancy was ignored. Sexual problems, value dilemmas and other special concerns of students were not investigated.

### *School Phobia*

Recent advances in treating school phobia appear positively related to accurate diagnosis and careful history-taking. In an eight-year study, Kennedy (1965) found that fifty school phobic cases responded to treatment with complete elimination of school phobia. The therapy described by Kennedy involved: 1) quick referral, 2) forced school attendance, 3) brief structured interviews with parents and with the child, and 4) specific instructions to parents and school personnel on child management. A careful selection of subjects for this treatment was no doubt a factor in its success.

School phobics are almost always treated individually with a specific treatment. Garvey and Hegrenes (1966) demonstrated that systematic desensitization as a treatment of choice proved successful. For twenty consecutive days, including Saturdays and Sundays, a therapist worked for twenty to forty minutes with a seven-year old on a step-by-step desensitization procedure carried out entirely in the school environment. Getting out of a car at school, going to the steps, walking up the steps, entering school and approaching the classroom constituted some of the minute stages of therapy. The entire process was clearly outlined by the authors. No subsequent manifestations of the phobia have appeared after two years.

The traditional approach to school phobia, typified by differential diagnosis and individual treatment, may ultimately serve as a model for counseling students with other special problems.

### *Behavior Problems and Delinquency*

Systematic exposure to identification models who exhibit socially appropriate behavior can have a positive influence in changing the behavior of the juvenile offender. Three recent studies in counseling with delinquents illustrated variations of social modeling techniques as treatment or as an aspect of treatment.

Persons and Pepinsky (1966) treated 41 boys at a reformatory matched with controls on age, intelligence, race, social class background, number of arrests, type of offense, length of incarceration to date and current adjustment. For twenty weeks, the experimental boys met twice weekly for one

and one-half hours of group therapy plus one hour of individual therapy. After each session, each boy anonymously identified another boy in his group as the peer leader. Pre- and posttreatment measures taken on therapists and subjects included the Taylor Manifest Anxiety Scale, Delinquency Scale, Minnesota Multiphasic Personality Inventory and the Kuder Preference Record.

The therapy, described as encouraging warm interpersonal relationships with each boy, developing an understanding and supportive atmosphere and developing insight into behavior, was inadequately outlined and not readily replicable. Thirty of the 41 experimentals (as compared to 12 of the 41 controls) were considered successfully treated on the basis of behavioral observations and test scores. Test scores for the thirty successfully treated boys moved more toward the scores of the therapists than toward those of the perceived peer leaders. The authors warned that such "convergence" may not be an inevitable concomitant of successful treatment, but may be a phenomenon correlated with the subjects judged to be improved.

Sarason (1968) reported pilot studies of counseling with institutionalized delinquents. Control and experimental groups were matched on age, intelligence, and severity and chronicity of delinquency. Experimental groups of four boys met for 15 sessions with two graduate students who served as group leaders-social models. Treatment consisted almost entirely of role-playing life situations of relevance to the boys, such as applying for a job, talking with a policeman and dealing with an angry father. In one experimental group the leader-models acted the roles; then the boys each role-played what had been modeled. In another experimental group, only the boys played the roles. Controls received no treatment. Measurements were taken with two self-rating forms of the semantic differential, Wahler's self description inventory, a cottage behavior rating scale, a weekly behavior summary, individual diagnoses and review board placement decisions. Preliminary analysis showed, in general, that the boys who received the modeled role-playing treatment changed the most in behaviors and attitudes. High anxiety and neuroticism were correlated with higher change in the modeling groups. Sarason is to be commended for his explicit and detailed account of the treatment procedures and for conducting a continuing sequence of experiments over time, rather than conducting an *ad hoc* study.

Truax et al. (1966) treated eight groups of ten patients each which met twice weekly for 24 one-hour sessions over three months. Four groups were hospitalized mental patients; four were juvenile delinquents. One half of the group received Vicarious Training Pretherapy (an audio tape of excerpts of "ideal" client behavior played for the group); the other half of the groups received 14 extra sessions, starting after the tenth regular session, with the therapist absent. The treatment, labeled "group therapy," was not described. Results of pre-post Q-sorts of positive and negative state-

ments showed the mean adjustment was in a negative direction for the juvenile delinquents. Effects of Vicarious Training Pretherapy treatment, however, suggested that the modeling tape facilitated change as measured by Q-sort. The use of Q-sorts leaves unanswered how and to what extent, if at all, the problem-related behaviors of the subjects changed.

Complementing the trend toward the use of modeling procedures are studies of individual treatment of problem behavior with behavioral techniques, illustrated by Wetzel (1966). One boy's compulsive stealing behavior was virtually eliminated by making a valued relationship with a friend contingent upon certain behavior. This study highlighted the importance of total staff involvement and participation in the success of any behavior modification undertaking.

### *Cultural and Ethnic Groups*

Research on counseling black students in public schools has begun to appear in the literature. Gilliland (1968) found that black adolescents who were provided small group counseling significantly increased their scores on the Cooperative English Achievement Tests, Occupational Aspiration Scale, and Vocational Development Inventory; they also improved their Grade Point Average, compared to nontreated controls. Two experimental groups, one of seven boys and another of seven girls, received something called "group-centered counseling" once a week for the academic year. The results seemed altogether too sensational to be attributed solely to "counseling." What actually occurred in the weekly sessions? Impressive findings like these merit attention and replication, but the treatment was so vaguely described that replication would prove to be most difficult.

Schaeffer and VonNessen (1968) described a non-experimental study in which group counseling and crisis-event role-playing were used with acting-out black adolescent girls. After four group sessions and a few individual contacts, the girls, according to the authors, developed skills in handling aggression, had better relationships with peers and teachers, and were not in any trouble requiring discipline. Unfortunately, no objective measures, controls or procedures of analysis were reported.

Thoresen (1967) used a behavioral approach in counseling one disadvantaged black youth who was not considered college material. During the student's first year at a junior college, he participated in weekly individual counseling (involving selective verbal and non-verbal reinforcement and modeled role-playing), weekly individual tutoring in reading and writing, and had part-time employment as an assistant in a laboratory; he was paid one dollar per hour for attending class, taking notes and being tutored. At the end of his first year the student had a "low B average" in transfer level courses. This case study does not, of course, demonstrate exactly what combination of activities may have promoted academic success. Experimentally designed, longitudinal studies could answer some important questions raised by this case study.



These beginnings, while inadequate, represent a significant step in research into areas such as counseling black students, where counselors need to know more about what to do and how to do it. Studies of counseling with ethnic and cultural minorities will undoubtedly increase dramatically during the next few years.

*Additional References:* Blakeman (1967); Day (1967); Duncan (1965); Hosford (1968); Kuntz (1966); Laxer et al. (1967); Mason (1968); Stewart and Moulton (1966).

### *Special Educational Problems*

Concern with the academic achievement of students has produced a deluge of writing and research. Since counseling students with special educational problems has historically been the forte of the counselor, it is not surprising that the bulk of material found for review in this chapter falls into this section. During the last three years, the topic of underachievement was the most popular special problem studied (19 of the 39 research articles reviewed were on underachievement). In contrast, little interest was displayed in counseling students with special educational problems involving academic attitudes, motivation, aspiration levels and decision-making. The inclusive characteristic of underachievement as a construct has interfered with clearly defining other educational problems.

#### *General Academic Achievement*

Kramer (1968) in working with college students successfully increased the oral participation behaviors of a group of college students. Sixty freshmen in a study skills course were assigned to six groups, three experimental and three control. Each group, led by a male counselor, had three men and seven women in it and met for six one-hour sessions. In reinforcement counseling, the counselor verbally reinforced responses termed Questioning, Responsibility and Positive; "traditional counseling" control groups emphasized reflection, clarification and interpretation. All groups listened to a social model tape demonstrating desired responses during the first and fourth interview. The reinforcement group significantly increased their participation behaviors compared to the traditional group, although most of the increase was accounted for by one of the three treatment counselors. This finding pointed out that counselor responses intended to be reinforcing stimuli were not equally effective. Future investigations are needed to explore the differential effects of counselor and client characteristics as well as the effects of modeling and expectancy to discover why some individuals change more than others. Growth curves which plot ongoing changes of individual verbal responses would provide important data on such questions.

Brown (1965), using scores on the Survey of Study Habits and Atti-



tudes, the Effective Study Test and Grade Point Average, showed that peer counselors were effective with large groups of college freshmen. Unfortunately, the treatment was not compared to a control or some other treatment, nor was information presented on what happened in group counseling. Despite the serious inadequacies of Brown's report, using trained peers as counselors deserves further investigation.

### *Underachievement*

The notion of underachievement is very confused and complicated. The causes and characteristics of underachievement are suggested by a wealth of labels such as free-floating anxiety, negative self-value, hostility toward authority, high dependence-independence conflict and negative interpersonal relations (Taylor, 1964; Thelen and Harris, 1968). Underachievement is usually defined arbitrarily by the investigator. As a result, depending on definition, underachievement is found both among gifted students (Ewing and Gilbert, 1967) and low-ability students (LeMay and Weigel, 1966; McGowan, 1968).

### *Subject Variability*

Different methods of selecting subjects not only yield different types of underachievers, but also significantly influence outcomes. The study by Winkler et al. (1965) provided an example of negative findings due to improper subject selection and poor research design. A group of 121 heterogeneous fourth-graders, defined as underachievers by a derived GPA and WISC Verbal Scale IQ, participated in one of five experimental conditions. No differences were found, probably due in part to individual differences. Some students may have had reading problems; others may have needed individual personal attention; others may have been motivated but lacked certain academic skills. Offering a treatment, e.g., "group counseling," to a mixed group is highly likely to produce no average gain, since gains by some subjects are cancelled by losses of others. The Winkler study characterizes much that is published in counseling and offers little promise of advancing professional knowledge. Such research should be discouraged.

Dickenson and Truax (1966), in contrast, limited the population studied to a certain type of underachiever, college freshmen whose ACT-predicted GPA was 2.2 or higher, but whose first semester grades were between 1.49 and 2.00 ( $C = 2.00$ ). Of 109 students, 48 accepted an invitation for group therapy, thus further homogenizing the subjects on motivation. Students were randomly assigned to treatment conditions. Compared to the controls, striking GPA improvements were found for the treated groups after 24 counseling sessions. Homogeneity on relevant factors may have contributed to the positive finding. This study would have been a greater contribution if the nature of the treatment had been

specified, permitting replication, and if the experimenters had controlled for the Hawthorne Effect.

Benson and Blocher (1967) also selected subjects from a homogeneous population, tenth-grade underachieving boys with negative feelings and attitudes toward school. The 28 boys who agreed to participate were randomly assigned to experimental and control groups. After 18 weeks of group counseling, thoroughly described but difficult to replicate, the experimental groups showed GPA improvement that was significantly greater on the average than the controls.

Chestnut (1965) suggested that since data on underachievers may not be linear, different treatments may not produce changes within subjects within the same amount of time. In comparing two types of group counseling for underachieving male college students, Chestnut found that one group had significantly greater rates of change in GPA after counseling than the other counseling group or the controls had. Gilbreath (1967), working with the same students as Chestnut, discovered that treatment effects were related to personality variables as assessed by the Sterns Activity Index. Those subjects whose scores indicated strong dependency needs, guardedness, emotional constriction, submission, orderliness and deliberation were more likely to improve in GPA if they participated in counselor-structured group counseling. The converse held true for their polar opposites who were more likely to improve in GPA in group-structured counseling. Since assignment to treatment groups was not based on Sterns Activity Index scores, both personality types existed in all groups, a fact (among many) which may have been partially responsible for neutralizing the effects of the two treatments.

LeMay and Weigel (1966) looked for possible differential effects in group counseling by focusing on study skills with high- and low-ability groups of poorly achieving college freshmen. At the end of one term, the high-ability experimental groups had a significantly higher GPA than the low-ability groups and all the controls, although the low-ability experimental groups had a higher GPA than the low-ability controls only. The effects of treatment over three months were known only for the high-ability experimentals, who maintained a significantly higher GPA than the high-ability controls.

### *Specificity of Treatment and Criterion*

The nature of the "treatment" is often insufficiently outlined in the reports of experiments. Authors and editors, rather than limiting treatment descriptions to a few sentences or labels, should insist on accurate, extensive and detailed treatment descriptions including observation schedules and check lists of activities. The profession should find no comfort in the widespread notion that one counseling treatment equals another counseling treatment with the same name or that something labeled "counseling"

applies similarly across subjects. Ewing and Gilbert (1967, p. 236) typified, unfortunately, the usual treatment description in stating:

The nature of the counseling received can be summarized as a combination of normal counseling procedures, in which the counselor attempts to be helpful to the client in regard to whatever problems the client presents, and counseling aimed especially at assisting students achieve better grades.

However, what actually are "normal procedures" and "helpful attempts?"

Katahn et al. (1966), illustrating the trend toward differential treatment, used systematic desensitization coupled with suggestions and advice to assist 14 test-anxious, underachieving college students develop skills for improved academic performance. In a thorough report of the treatment, Katahn outlined the eight, one-hour sessions which resulted in significant increases in experimental group GPA and lower anxiety scores, compared to the controls. Using highly motivated volunteers with specific objectives in mind, who are exposed to specific brief treatment relevant to their psychodynamic needs, resulted in positive outcomes.

Thoresen and Neuman (1968) found that group desensitization methods in general caused significantly greater decreases in mean change scores than group insight procedures. Two professionals and two sub-professionals (first semester graduate students) each conducted both treatments to group of three for five sessions over a five-week period. Sub-professional counselors were, in general, as effective as the professionals. Both treatment groups showed significant decreases on self-report measures, an observer checklist of anxious behaviors and physiological data when compared to wait controls and no-contact controls. A one-year follow-up is presently being completed.

Roth et al. (1967) concluded that control and specificity of therapeutic approaches could lead to developing more explicit and effective counseling techniques, which could then be applied differentially to certain kinds of underachievers.

Improved GPA is a commonly found short-term gain in many underachievement investigations (Benson and Blocher, 1967; Schmieding, 1966; Dickenson and Truax, 1966; Thelen and Harris, 1968; Katahn et al., 1966; Leib and Snyder, 1967). Many contradictory and negative findings are also common (Heller and Gurney, 1968; Hill and Grieneeks, 1966; Chestnut, 1965). Long-range results need greater attention, but unfortunately, few long-range follow-ups have been published. Goodstein (1967), in following the original study by Marx (1959), found that initial gain in GPA was completely negated five years later. In fact, higher proportions of control subjects graduated than did counseled subjects.

The rationales for length, intensity, frequency and schedule of treatment are also important concerns that are seldom discussed. For example, Schmieding (1966) exposed his subjects to three counseling sessions of

one-half hour in length over one semester. Why three? At the other extreme, Dickenson and Truax (1966) held weekly one-hour sessions for 12 weeks. Why 24 sessions?

*Additional References:* Abel (1967); D'Zurilla (1966); Hendrix (1965); Katahn (1967); LeMay (1966); Peres (1965); Shepherd (1965).

### *Implications and Conclusions*

This selective review of research on counseling students with special problems revealed four major trends: 1) a striking increase in the use of group counseling; 2) increased concern for behavior change outside the interview and its relationship to treatment; 3) increased specificity of behavior to be changed; and 4) some use of differential and specific treatment.

The conspicuous absence of research activity on counseling students who have neurological, physical and communication disabilities was disturbing. Special problems confronting the very slow or the very rapid learner also received little attention (Ramsey, 1967; Sulzbacher and Houser, 1968; Birnbrauer et al., 1965; Wiesen and Watson, 1967; Drews, 1965).

To deal more effectively with the complex human concerns of students with special problems, the profession needs well-designed experimental studies, others that replicate and follow-up prior well-designed studies, and some large-scale cooperative studies.

Investigators should consider specific, psychodynamic and behavior-related diagnosis before treatment. Specific delineation of treatment is vital to the achievement of an understanding of what works. A variety of criterion measures must be considered as well as new ways of analyzing data. Dickenson and Truax (1966), in examining pre- and posttreatment results, counted the number of subjects who moved from probation to nonprobation status. Such data may be more important than determining whether or not significant increases in average GPA occurred. Helping one college student raise his GPA from 1.96 to 2.01 may be, for him, the difference between graduating or not graduating. In the probation-or-nonprobation, pass-or-fail, graduate-or-dropout realistic world of the student, the idea of seeking practical and statistical solutions to problems seems eminently rational.

Research is needed which comes closer to the single-subject, own-control design that has long characterized operant conditioning studies. (See the *Journal of Applied Experimental Analysis*.) Such a research model would focus primary attention on the behavior of individuals, baseline, treatment and follow-up performances, and would move away from the increasingly limited value of designs requiring large *N*'s concerned only with mean differences.



## Bibliography

- Benson, R. L. and Blocher, D. H. "Evaluation of Developmental Counseling with Groups of Low Achievers in a High School Setting." *School Counselor* 14: 215-220; 1967.
- Birnbrauer, J. S. et al. "Classroom Behavior of Retarded Pupils with Token Reinforcement." *Journal of Experimental Child Psychology* 2: 219-235; 1965.
- Brown, W. F. "Student to Student Counseling for Academic Adjustment." *Personnel and Guidance Journal* 43: 811-817; 1965.
- Chestnut, W. J. "The Effects of Structured and Unstructured Group Counseling on Male College Students' Underachievement." *Journal of Counseling Psychology* 12: 388-398; 1965.
- Dickenson, W. A. and Truax, C. B. "Group Counseling with College Underachievers." *Personnel and Guidance Journal* 45: 243-247; 1966.
- Drews, Elizabeth M. "Counseling for Self-Actualization in Gifted Girls and Young Women." *Journal of Counseling Psychology* 12: 167-175; 1965.
- Ewing, T. N., and Gilbert, W. M. "Controlled Study of the Effects of Counseling on the Scholastic Achievements of Students of Superior Ability." *Journal of Counseling Psychology* 14: 235-239; 1967.
- Garvey, W. P. and Hegrenes, J. R. "Desensitization Techniques in the Treatment of School Phobias." *American Journal of Orthopsychiatry* 36: 147-152; 1966.
- Gilbreath, S. H. "Group Counseling, Dependence, and College Male Underachievement." *Journal of Counseling Psychology* 14: 449-453; 1967.
- Gilliland, B. E. "Small Group Counseling with Negro Adolescents in a Public High School." *Journal of Counseling Psychology* 15: 147-152; 1968.
- Goodstein, L. D. "Five-year Follow-up of Counseling Effectiveness with Probationary College Students." *Journal of Counseling Psychology* 14: 436-439; 1967.
- Heller, Blanche and Gurney, D. "Involving Parents in Group Counseling with Junior High Underachievers." *School Counselor* 15: 394-397; 1968.
- Hill, A. H. and Grieneeks, Laurabeth. "An Evaluation of Academic Counseling of Under- and Over-Achievers." *Journal of Counseling Psychology* 13: 325-328; 1966.
- Katahn, M.; Strenger, S.; and Cherry, Nancy. "Group Counseling and Behavior Therapy with Test-Anxious College Students." *Journal of Consulting Psychology* 30: 544-549; 1966.
- Kennedy, W. A. "School Phobia: Rapid Treatment of Fifty Cases." *Journal of Abnormal Psychology* 70: 285-289; 1965.
- Kramer, H. C. "Effects of Conditioning Several Responses in a Group Setting." *Journal of Counseling Psychology* 15: 58-62; 1968.
- Leib, J. W. and Snyder, W. U. "Effects of Group Discussion on Underachievement and Self-Actualization." *Journal of Counseling Psychology* 14: 282-285; 1967.
- LeMay, M. L. and Weigel, R. G. "Group Counseling with High and Low Ability College Freshmen." *Journal of Educational Research* 59: 429; 1966.
- Marx, G. L. A Comparison of the Effectiveness of Two Methods of Counseling with Academic Underachievers. Doctor's thesis. Iowa City: Univ. of Iowa, 1959. Abstract: *Dissertation Abstracts* 20: 2144-45; No. 6, 1959.
- McGowan, R. J. "The Effect of 'Brief Contact' Interviews with Low Ability, Low Achieving Students." *School Counselor* 15: 386-389; 1968.
- Persons, R. W. and Pepinsky, H. B. "Convergence in Psychotherapy with Delinquent Boys." *Journal of Counseling Psychology* 13: 329-334; 1966.
- Ramsey, G. V. "Review of Group Methods with Parents of the Mentally Retarded." *American Journal of Mental Deficiency* 71: 857-863; 1967.
- Roth, R. M.; Mauksch, H. O.; and Peiser, K. "The Nonachievement Syndrome, Group Therapy, and Achievement Change." *Personnel and Guidance Journal* 46: 393-398; 1967.
- Sarason, I. G. "Verbal Learning, Modeling, and Juvenile Delinquency." *American Psychologist* 23: 254-266; 1968.
- Schaeffer, D. T. and Von Nessen, R. "Intervention for Disadvantaged Girls: Insight for School Faculties." *American Journal of Orthopsychiatry* 38: 666-671; 1968.
- Schmieding, O. A. "An Investigation of Efficacy of Counseling and Guidance Procedures with Failing Junior High School Students." *School Counselor* 14: 74-80; 1966.
- Sulzbacher, S. I. and Houser, Joyce E. "A Tactic to Eliminate Disruptive Behaviors in

- the Classroom: Group Contingent Consequences." *American Journal of Mental Deficiency* 73:88-90; 1968.
- Taylor, R. C. "Personality Traits and Discrepant Achievement: a Review." *Journal of Counseling Psychology* 11: 76-82; 1964.
- Thelen, M. H. and Harris, C. S. "Personality of College Underachievers Who Improve with Group Psychotherapy." *Personnel and Guidance Journal* 46: 561-566; 1968.
- Thoresen, Carl E. "A Behavioral Approach to Encouraging College Accomplishment in Disadvantaged Youth: An Exploratory Study." *Journal of College Student Personnel* 8: 173-175; 1967.
- Thoresen, Carl E. and Neuman, Donald R. "Professional and Subprofessional Counselors Using Group Desensitization and Insight Procedures to Reduce Examination Anxiety." Stanford, California: Stanford University, School of Education, 1968. (Mimeo.)
- Truax, C. B. et al. "Changes in Self-Concepts During Group Psychotherapy as a Function of Alternate Sessions and Vicarious Therapy Pretraining in Institutionalized Mental Patients and Juvenile Delinquents." *Journal of Consulting Psychology* 30: 309-314; 1966.
- Wetzel, R. "Use of Behavioral Technique in a Case of Compulsive Stealing." *Journal of Consulting Psychology* 30: 367-374; 1966.
- Wiesen, A. E. and Watson, Edith. "Elimination of Attention-Seeking Behavior in a Retarded Child." *American Journal of Mental Deficiency* 72: 50-52; 1967.
- Winkler, R. C. et al. "The Effects of Selected Counseling and Remedial Technique on Underachieving Elementary School Students." *Journal of Counseling Psychology* 12: 384-387; 1965.

### Additional References

- Abel, W. H. "Group Counseling and Academic Rehabilitation of Probationary Transfer Students." *Journal of College Student Personnel* 8: 185-188; 1967.
- Blakeman, J. D. *The Effects of Activity Group Counseling on the Self-Evaluation and Classroom Behavior of Adolescent Behavior Problem Boys*. Doctor's thesis. Athens: University of Georgia, 1967. Abstract: *Dissertation Abstracts* 28:2066; No. 6, 1968.
- Day, S. R. *The Effects of Activity Group Counseling on Selected Behavior Characteristics of Culturally Disadvantaged Negro Boys*. Doctor's thesis. Athens: University of Georgia, 1967. Abstract: *Dissertation Abstracts* 28: 3969; No. 10, 1968.
- Duncan, J. A. *The Effects of Short-term Group Counseling on Selected Characteristics of Culturally Deprived Ninth Grade Students*. Doctor's thesis. Athens: University of Georgia, 1965. Abstract: *Dissertation Abstracts* 27: 387-388; No. 2, 1967.
- D'Zurilla, T. J. "Persuasion and Praise as Technique for Modifying Verbal Behavior in a 'Real-Life' Group Setting." *Journal of Abnormal Psychology* 71: 369-376; 1966.
- Hendrix, O. R. "The Effect of Special Advising on Achievement of Freshmen with Low Predicted Grades." *Personnel and Guidance Journal* 44: 185-188; 1965.
- Hosford, Ray E. *Devising Social-Model Counseling Procedures for Elementary School Children*. Washington, D. C.: American Educational Research Association, 1968. (ED 017 029, MF \$0.25, HC \$0.60).
- Katahn, M. "Systematic Desensitization and Counseling for Anxiety in a College Basketball Player." *Journal of Special Education* 1: 309-314; 1967.
- Kuntz, A. B. *An Experimental Evaluation of Short-Term Group Counseling with Non-conforming Adolescents*. Doctor's thesis. Fargo: University of North Dakota, 1966. Abstract: *Dissertation Abstracts* 27: 3290-A; No. 10, 1967.
- Laxer, R. M.; Quarter, J. J.; Isnor, Catherine; and Kennedy, D. R. "Counseling Small Groups of Behavior Problem Students in Junior High Schools." *Journal of Counseling Psychology* 14: 454-457; 1967.
- LeMay, M. L. *An Experimentally Controlled Investigation of the Effects of Group Counseling with College Underachievers*. Doctor's thesis. Eugene: University of Oregon, 1966. Abstract: *Dissertation Abstracts* 27: 3694-3695-A; No. 11, 1967.
- Mason, Evelyn P. "Project Catchup: An Educational Program for Junior High Students from American-Indian, Mexican and Caucasian Ethnic Backgrounds." *Psychology in the Schools* 5: 272-276; 1968.

- Preus, J. B. *The Effect of Four Student Personnel Services on the Academic Performance of Underachieving Arts College Freshmen*. Doctor's thesis. Minneapolis: Univ. of Minnesota, 1964. Abstract: *Dissertation Abstracts* 25: 4549-50; No. 8, 1965.
- Shepherd, R. E. "The Relation of Counseling and Student Problems to Graduation." *Journal of Counseling Psychology* 12: 244-247; 1965.
- Stewart, L. H. and Moulton, R. W. *Increasing the Academic Achievement of Culturally Disadvantaged Youth*. Berkeley: Univ. of Calif., 1966. (ED 013 144, MF \$0.50, HC \$3.60).

## AUTHOR

Island, D. David Address: University of Washington at Seattle, Washington Title: Assistant Professor Age: 33 Degrees: B.S., Portland State College; M.A., Univ. of Minn.; Ph.D., Univ. of Minn. Specialization: Nonverbal Behavior in Counseling.

## 9: SYSTEMS APPROACHES IN GUIDANCE

William W. Cooley\*

*University of Pittsburgh*

Raymond C. Hummel

*University of Pittsburgh*

This is the first edition of the *Review of Educational Research* to include a chapter on systems approaches in guidance. In this new area, problems with terms and definitions abound. A definition for this chapter was educed from the "systems" literature and from papers and project reports furnished by colleagues who are developing systems for guidance. Because the topic is new, the following format was followed: 1) an overview of the systems approach and how it differs in perspective from traditional research, 2) a review of three systems approaches in guidance, 3) a brief summary of a few other efforts in this general area, and 4) projections for the future.

### *The Systems Approach*

A systems approach is a method for analyzing and realizing the values, goals or policies of a human enterprise. The method makes explicit "the structure or organization of an orderly whole, clearly showing the inter-

---

\*Dr. David Tiedeman, Harvard University, and Dr. Donald E. Super, Teachers College at Columbia, served as consultants to Drs. Cooley and Hummel on the preparation of this chapter.



relationships of the parts to each other and to the whole itself" (Silvern, 1965, p. 2). At least three steps are involved in developing a systems approach:

1. Translate the broad aims of the enterprise into objectives which are explicit and operational.
2. Design the procedures which are intended to accomplish these objectives, identify the relevant variables which the procedures are intended to order or change, and construct a model which suggests *a priori* and consequent relationships among the identified variables.
3. Implement the model and evaluate the results of the innovation in terms of the operationally stated objectives.

A systems model thus includes specific statements about objectives, procedures, information requirements and the activities of each participant. The evaluation activities are ordinarily intended not only to yield judgments of the worth of what has been done but also to provide "feedback," i.e., information useful for revising the original objectives and procedures. Thus, a system in operation often has a cybernetic character; the model keeps changing as a result of information obtained in its implementation.

Discussions of the systems approach are often confounded by the systems language, which tends to present rather simple ideas in complicated ways. Terms such as *interface*, *mission-oriented*, *algorithms*, *simulation*, *feedback*, *cybernetic loops*, and *deviation-counteracting processes* are defined concepts within a systems theory. However, they tend to confuse the outsider either because of their unfamiliarity or because the "household meanings" of some systems concepts differ from their technical usage. The interested reader who would like to become more familiar with this general field should consult Silvern (1965) or Pfeiffer (1968).

Ordinarily, systems applications involve the use of machines and the deliberate coordination of the activities of men and the operations of machines. One job of a systems analyst is to determine which tasks are best done by men and which are best done by machines. The computer, of course, is the machine which has most often been used. The introduction of machines requires a more systematic and explicit definition of program objectives. Gagné (1962) brought together several important papers that are required reading in this area. Wiener (1950) provided a classic description of cybernetics which would serve "the human use of human beings." More specific to the computer-counselor problem is a recent paper by Super (1968a) in which he considered how the functions of the counselor will change as guidance computer systems are developed.

### *Systems Approach and Research*

The systems approach extends beyond the traditional research emphasis on theory-building into the domain of development. A systems approach involves the application of the methods of science to problems of the

engineering and development of practical programs. The increasing emphasis on development, an important new trend, is likely to have a greater impact on guidance programs and practices than more conventional educational research has had. Traditional distinctions between research and development are likely to disappear as they are found to handicap the improvement of guidance practice.

The systems approach also differs from traditional research because it is almost necessarily a team effort. Published reports of isolated studies by lone investigators have had little impact on the practice of guidance. Research and development team efforts promise to produce a more reliable translation of research findings into usable educational programs. Cooperative enterprises require larger budgets and more complete research management procedures than have been common in the past. However, funding sources have become increasingly willing to support such projects, and investigators have emerged who have the will and competence to manage them.

Large budgets and team efforts in themselves do not guarantee successful implementation of research findings. Project TALENT (Flanagan et al., 1962), for example, had both, and although the results of Project TALENT continue to contribute to the research literature, they have had a negligible impact upon guidance practice. The research reports from TALENT, as well as those from most research projects, have not modified guidance practices because they have not been accompanied by 1) specific objectives for guidance, 2) identification of the necessary specific information, 3) definition of a set of procedures that schools might use to meet guidance objectives, 4) tests and evaluation of those procedures, 5) packaging of the results (tests, manuals, computer procedures, etc.), and 6) dissemination of developed programs to the schools. An approach which corrected each of these failures would exemplify a systems approach to research and development in guidance and would be more likely to produce desired innovations than traditional modes of research produce.

The concept of "system" is based on order or interdependence among phenomena (Hills, 1967). One is attempting to locate or create an order among guidance functions in using a systems approach. In doing so, it is necessary to decide how extensive is the range of phenomena or functions one wishes to order. Every system can be hierarchically viewed as a subsystem within a larger system. A delimited set of guidance procedures for processing and communicating information to a student for self-appraisal might be systematized. All the guidance functions in an educational institution might be coordinated within one system involving students, school personnel and facilities in organized cycles of information-processing, appraisal, decision-making and planning. Such a systematized guidance program might also be interrelated with other subsystems to form an educational system. General systems theory would already construe a

school as a system; the major concern in this chapter, however, is the specific application of systems theory to a deliberate strategy for defining and ordering certain educational and guidance functions.

From a hierarchical perspective, all the systems reviewed in this chapter appear to be subsystems of (partial) guidance functions. The major systems projects in guidance are concerned mainly with computer-assisted occupational exploration; they make sophisticated use of the computer's capacity for providing random access to data. To date, no attempt has been made to systematize all the guidance functions in an educational program, although such an aim is implied in Flanagan's (1967) prospectus for Project PLAN.

### *Systems Projects*

The three projects described in this section illustrate major contemporary efforts to apply systems techniques to guidance. Each project was concerned with career guidance and involved using computers.

#### *The SDC Vocational Counseling System: Autocoun*

A computer-based educational and vocational counseling system was developed at the System Development Corporation (SDC) by Cogswell and Estevan (1965) as part of a larger project to develop procedures for systems analysis and computer simulation in education. The SDC Vocational Counseling System was designed to automate certain information-processing functions ordinarily carried out by the counselor. The system is focused on providing vocational and educational information to students who interact directly with a time-sharing computer. The counselor can also program special operations and data for specific students.

*Procedures.* Autocoun (Loughary, Friesen and Hurst, 1966) uses a conversational mode of computer language which does not require that users have computer programming skills. The autocoun operation simulates the counselor's assessment of information about a student prior to and during an interview. Information on a population of students in a particular school is stored in the computer together with information about the world of work and the local and state post-high school training institutions. The information includes grades, test scores, teacher comments and other data from school records. Similar information and follow-up data are included on former students. The information on post-high school training describes kinds of institutions, entrance requirements, fields of study and scholarship opportunities.

The computer is programed to provide on-line and off-line retrieval of student information and to generate special purpose documents such as report cards and lists of failing students. The program can track and monitor student progress and use current and follow-up student data to



generate multiple regression formulas for predicting student performance. Using a teletype keyboard, the student asks for relevant information from the computer data files. The student is asked to supply information about his difficulties with present courses, future course choices and post-high school plans. The computer then prepares off-line reports, based on the stored information, advising each student on the probable outcomes of his tentative choices. The computer also prepares a summary of the student's on-line interview for the counselor. Plans call for programing the computer to conduct three on-line interviews, one each for high school program planning, post-high school educational planning and vocational planning.

*Evaluation.* An initial evaluation of the automated system was made on a sample of forty ninth-grade students randomly selected from a junior high school population in Palo Alto, California (Loughary, Friesen and Hurst, 1966). No significant differences were found between two school counselors and the computer simulations on three-fourths of the appraisal statements completed by the students. For example, the human counselors and the computer performed similarly in identifying pupils who were underachieving and pupils who had overly ambitious plans. The automated counseling system predicted significantly higher grade point averages, predicted more potential student dropouts, and encouraged students to explore a wider range of academic subjects than did the counselors.

The students were generally in favor of using Autocoun. They thought it had more specific and factual information about college requirements than did the counselor, and they communicated with Autocoun about a significantly greater number of course-related problems. About half of the students thought Autocoun did not give enough consideration to personal interests and personality characteristics. All but two of the students favored using Autocoun in the school, but they wanted it used in conjunction with the counselor.

Understandably, the SDC investigators found the simulation of the counselor's logical appraisal procedures easier to achieve than the automation of interviewing procedures. There are subtle but important individual differences among students in the way they feel about and learn to interact with the machine for self-exploration (Cogswell, 1966). Some students may find the need for human contact increased by using a machine; for other students the need for contact may be reduced.

### *Information System for Vocational Decisions*

The most ambitious system-building effort in guidance is the Information System for Vocational Decisions (ISVD) being developed at Harvard by Tiedeman and his colleagues. Of the vast number of technical reports and working papers which have been produced to date, the annual reports for 1967 and 1968 appear to be the most useful and most readable for



reviewing the work of ISVD (Tiedeman et al., 1967, 1968). The developers characterize the ISVD as follows:

The major objective of ISVD is to improve vocational decision-making through the use of a computer guidance system. The program is to be so designed that the student can relate knowledge about himself to data about education, training, and work, and thereby create a body of information upon which he can base his career decision. The entire program links person, computer, and teacher or counselor in such a way that the student can conduct a dialogue with the computer while the counselor assists in interpreting and evaluating the results of the dialogue. (Tiedeman et al., 1968, p. 1)

Sub-objectives which may be drawn from this major objective are: 1) to teach decision-making, 2) to provide information needed by the decision-makers, and 3) to develop within the decision-maker an awareness that he is able, to a considerable extent, to determine the course of his own career (i.e., to develop "a sense of agency"). The scope of the information base is extensive since the objective of ISVD is to provide career guidance to all. Thus, the information base includes all levels of education and training, military service, and all aspects of occupations including the latest "help wanted" data.

*Procedures.* Central to the ISVD system is the concept (and the computer program) called MONITOR. It is MONITOR's job to develop decision-making skills and attitudes while the student interacts with the data. Simulation games are also presented by the computer. In initial encounters with the system the student is encouraged to simulate his career in a game similar to the Life Career Game (Boocock, 1968), in which alternative decisions are made and the student "lives through" the consequences. It is assumed that simulated experiences will better enable the student to use the system "in the reality of his own life."

Defining all the objectives specifically for design and implementation of this system constitutes a major project. The extensive set of operational definitions produced should be generally useful in the development of guidance systems, even if a major objective of ISVD (i.e., a computer programmed to "converse" with the counselee in "natural English") is not fully realized.

Reservations about the natural language program objective arise from the high costs of programing for a conversational mode, especially when no format or content constraints are to be imposed upon the inquirer. A natural language approach may be essential to the achievement of ISVD objectives, but this might have been demonstrated first with a small segment of the career guidance problem, before undertaking a more massive development effort. If evaluation demonstrates that interaction with a natural language program produces more effective decision-making and a

greater "sense of agency" than other procedures, the builders of the ISVD will have to wait for computer scientists and engineers to reduce costs sharply before they will see widespread use of their system.

### *The IBM Guidance Counseling Support System*

The IBM Guidance Counseling Support System, reflecting the theories of Super (1968b), is designed to foster "vocational maturity" by helping the student deal with certain tasks which confront him at later stages of his vocational development. The system provides the student with experiences which 1) broaden his knowledge of work and his personal and occupational multi-potentiality, 2) provide information on local employment and job training opportunities in occupational areas of interest to him, 3) help explore college-relevant curriculum preferences exclusive of occupational goals and then relate preferences to occupational potentials, and 4) help narrow his search for post-high school training institutions which satisfy his curriculum preferences, career goals and personal preferences (Minor, 1968).

*Procedures.* In using the system, the student sits at a "learning station" in which information filmstrips are displayed under computer control. The student moves through the sequence by using a numeric keyboard to respond to multiple-choice questions. An attached typewriter-printer produces output describing the student's interaction with the computer.

The system contains information files on occupations, the military, universities and a local occupation training file. The computer also contains a personalized file summarizing the student's grades, academic aptitude and vocational interests.

*Evaluation.* Results from student use of this system are not yet available. A field test of the system in a New Jersey school is planned for the academic year 1968-1969. The results of the field test will be an important guide to further developments at IBM and elsewhere. Until the evaluation stage is reached and trial results begin to provide feedback for system improvement, the greatest potential for the systems approach cannot be realized.

### *Other Systems Projects*

The Computer-Assisted Career Exploration (CACE) System is an entirely computer-based guidance system developed at Pennsylvania State University under the direction of Impellitteri (1968). This system is intended to provide individually tailored occupational information for ninth-grade boys interested in vocational and technical courses of study. The objective of the system is to increase the student's knowledge of specific occupations and relevant occupational exploratory behavior. Although 80% of his subjects said that they were definitely aided by the

CACE experience, Impelliteri found no initial evidence of effects on their curriculum choices or occupational exploration.

Harris (1968) is currently developing an automated library of data at Willowbrook High School in Illinois for the use of counselors and students. The system includes a six-week orientation period for students in sophomore English classes. During the orientation period, students take a battery of tests and become familiar with Roe's two-dimensional occupational classification system. According to Harris, a primary objective of the student's interaction with the computer is to provide him with a model of decision-making applicable to adult life.

Magoon (1968; in press) developed a problem solving procedure for college students who are experiencing vocational-educational indecision; the system operates as a self-directed learning program. The student writes out responses to questions related to career planning. Meeting with groups composed of no more than five clients, the counselor comments on the data provided by each student and suggests where he might obtain relevant information. Although the program is designed to help the student make up alternative career plans, Magoon considered its essential use to be the learning of a problem-solving process for resolving career indecision.

In his System of Interactive Guidance and Information (SIGI), Katz (1966, 1968) utilized three subsystems of data involved in career decision-making: a value system, an information system, and a prediction system. Interacting with a computer, the student manipulates these data and generates a ranking of career options; these options can be revised as the three kinds of data change over time.

Hummel (1968) developed a Coordinated Information and Guidance System (CIGS) that provides a learning environment in which each student studies and is assisted in appraising himself, in clarifying his values and in obtaining relevant information. The basic operative unit is a group of 15 students which meets no more often than one school period daily and is typically led by a teacher or counselor; a "package" of exercises provides a permissive, loosely structured interaction among the students and the leader. The parts of the CIGS which prove their worth in school tryouts will be computer-based.

Simulation procedures involving the use of games and problem-solving exercises have been given increasing application in guidance systems. Krumboltz (1967) developed problem-solving experiences for five occupations, including accounting and laboratory technology. He found that the problem-solving "career kits" provided students with more than mere occupational information; the "career kits" consistently produced an increase in student interest and search behavior. Students from schools with a lower socioeconomic status responded more positively to the problem-solving approaches than students from middle-class schools.

A variety of social simulation games and their rationales were reported



by Boocock and Schild (1968). The Life Career Game (Varenhorst, 1968) appears to be especially adaptable as a subsystem within a guidance program. The Palo Alto School District has been using simulation procedures as part of a comprehensive program for fostering student decision-making.

An example of a more general system under development is the "multi-medium support system for educational and career planning" reported by Johnson, Youst and Burnham (1968). This system includes both a computer-based index and a "perceptual resources file" (mostly containing audio tapes and film slides). Special multiple life-career studies are being developed for each occupational area; each story will be illustrated in a five-minute presentation by controlled slide, movie and audio media. Counselors and students will be provided with film-making equipment for creating additional resource materials. The system is currently being tried out with eighth-grade students in an urban junior high school.

Project PLAN, a Program for Learning According to Needs, has been designed as a comprehensive individualized system of education (Flanagan, 1967). The student is given a series of units designed to develop his decision-making, planning and personal management skills. Relevant tests to assess the progress of the teacher, the counselor and the student are given at the completion of each teaching-learning unit. A basic objective of PLAN is to enable each student to create and revise his own program of educational and personal development during the years he is in school.

### *Looking Ahead*

Guidance "systems" have always existed. Certain aspects of guidance systems, especially the use of computers for vocational guidance, are currently being developed using something which approximates a systems analysis approach. However, the efforts reviewed in this chapter do not represent the full potential of systems analysis. All aspects of the guidance program are amenable to the systems approach. Thoresen (1968), for example, summarized the basic features and implications of the systems approach for counselor education.

Considering all guidance functions together, it becomes clear that one man's system is another man's subsystem. The systems approach soon requires looking at the guidance process within the context of education. Cooley (1968) argued that guidance systems should be considered as part of the total educational system, which itself is changing as a result of the new educational technology.

It seems appropriate to end this discussion with a note of enthusiasm and some words of caution. The new emphasis on systems approaches and the accompanying concern for research and development is a healthy sign. This approach will probably produce more powerful procedures for



assisting individuals to appraise their life prospects and to formulate relevant plans. It is also more likely that the systems approach will provoke more critical analysis of and innovations in guidance practices than have been produced by traditional research.

The first word of caution relates to philosophy. The practice of guidance in America has been firmly committed to individual freedom, initiative, and responsibility in personal planning and decision-making. Counseling, by definition, requires a human relationship in which the counselee feels free and responsible for the process and outcomes. A highly systematic, technologically-supported program of guidance services may appear to contradict such a commitment. All social and technological change is accompanied by the risk of depersonalization and the sacrifice of personal freedom to the demands of a system. The history of education demonstrates this risk all too persuasively. The problem, however, lies not in sophisticated technology, but in the leadership and organization of the social system that supports the technology. The systems development in guidance to date does not seem to have aggravated the problem of freedom. The guidance systems rather are aimed toward fostering individual information-seeking and decision-making. In fact, by providing the individual with more powerful media for exploring the meaning of information as it applies to him, and by permitting counselors and teachers to be the student's collaborators in interpreting the information, the new guidance systems may serve to expand rather than constrain personal freedom.

The final word of caution concerns the need to distinguish between tomorrow and someday. Myers (1968) discussed this concern and has pointed out the tendency for the developer to underestimate the time and money needed for his systems development effort in order to secure the necessary funds. When the actual long-term development needed becomes evident, it might result in certain disillusionment of those doing the funding and those who are eagerly awaiting the fruits of these efforts. Panaceas for guidance are no more possible tomorrow than they are today or than they were yesterday. It's a long, hard row to hoe, and the promise of the systems approach is to facilitate the hoeing.

### Bibliography

- Boocock, Sarane S. and Schild, E. O. (editors.) *Simulation Games in Learning*. Beverly Hills, Calif.: Sage, 1968.
- Cogswell, John F. and Estevan, D. P. *Explorations in Computer-Assisted Counseling*. SDC Document TM-2582/000/00. Santa Monica, Calif.: System Development Corporation, 1965.
- Cogswell, John F. *New Solutions to Implementing Instructional Media Through Analysis and Simulation of School Organization*. U. S. Office of Education, Final Report. Santa Monica, Calif.: System Development Corporation, 1966.
- Cooley, William W. *Computer Systems for Guidance*. Invited address to Division E, Amer. Educ. Res. Assoc. Annual Meeting, February 1968. Pittsburgh: Department of Educational Research, Univ. of Pittsburgh. (Mimeo.)

- Flanagan, John C. *Project PLAN: A Program of Individualized Planning and Individualized Instruction*. Paper presented at Project ARISTOTLE Symposium, Washington, D. C. Palo Alto, Calif.: American Institutes for Research, December 1967.
- Flanagan, John C. et al. *Design for a Study of American Youth*. Boston: Houghton Mifflin Co., 1962.
- Gagné, R. M. (editor). *Psychological Principles in System Development*. New York: Holt, Rinehart and Winston, 1962.
- Harris, Jo Ann. *Comments about Learning for Adult Activities and Roles Through Willowbrook's Computerized Vocational Information System*. DuPage County, Ill.: Willowbrook High School, 1968. (Mimeo.)
- Hills, R. Jean. *The Concept of System: An Analysis and an Illustration*. Paper presented to the Amer. Educ. Res. Assoc., February 1967. Eugene, Ore.: Center for the Advanced Study of Educational Administration, Univ. of Ore. (Mimeo.)
- Hummel, Raymond C. *A Coordinated Information and Guidance System*. Pittsburgh: Learning Research and Development Center, Univ. of Pittsburgh, 1968. (Mimeo.)
- Impellitteri, Joseph T. *The Development and Evaluation of a Pilot Computer-Assisted Occupational Guidance Program*. Pennsylvania Department of Public Instruction, Final Report, Project No. 18033. University Park: Vocational Educational Department, Penn. State Univ., July 1968.
- Johnson, Donald H.; Youst, D. B.; and Burnham, R. W. "The Application of Educational Media in a Multi-Medium Support System for Education and Career Planning." *Educational Media and Educational Development*. (Edited by Walter Lifton.) New York: John Wiley and Sons, 1968.
- Katz, Martin. "A Model of Guidance for Career Decision-Making." *Vocational Guidance Quarterly* 15: 2-10; 1966.
- Katz, Martin. *System of Interactive Guidance and Information (SIGI)*. Princeton, N. J.: Educational Testing Service, 1968. (Xerox.)
- Krumboltz, John D. *Vocational Problem-Solving Experiences for Stimulating Career Exploration and Interest*. Stanford, Calif.: Stanford Univ., August 1967. Abstract, pp. XX-XXV. (Mimeo.)
- Loughary, John W.; Friesen, Deloss; and Hurst, Robert. "Autocoun: A Computer-Based Automated Counseling Simulation System." *Personnel and Guidance Journal* 45: 6-15; September 1966.
- Magoon, Thomas M. *An Effective Problem Solving Model for Vocational-Educational Indecision Problems*. College Park, Md.: University of Maryland, 1968. (Mimeo.)
- Magoon, Thomas M. "Developing Skills for Educational and Vocational Problems." *Behavioral Counseling: Cases and Techniques*. (Edited by John D. Krumboltz and Carl E. Thoresen.) New York: Holt, Rinehart and Winston, in press.
- Minor, Frank J. *An Experimental Computer-Based Educational and Occupational Orientation System for Counseling*. Paper presented to XVth International Congress of International Association, Amsterdam, The Netherlands, August 1968. Yorktown Heights, N. Y.: IBM Advanced Systems Development Division. (Xerox.)
- Myers, Roger A. *Computer-Aided Counseling: Some Issues of Adoption and Use*. Paper presented to XVth. International Congress of the International Association of Applied Psychology, Amsterdam, The Netherlands, August 1968. New York: Teachers College, Columbia Univ. (Mimeo.)
- Pfeiffer, J. *A New Look at Education: System Analysis in Our Schools and Colleges*. New York: Odyssey Press, 1968.
- Silver, L. C. "Systems Analysis and Synthesis in Training and Education." *Automated Education Handbook*, November 1965, IC1-IC25.
- Super, Donald E. *Computers and Counselor Roles: Counselor Acceptance of Systems in Guidance and Eventual Functions of Counselors*. Paper presented at Symposium on the Use of Computers in Counseling and Guidance, National Vocational Guidance Association, April 1968. New York: Teachers College, Columbia Univ., 1968a. (Mimeo.)
- Super, Donald E. *Vocational Development Theory: Persons, Positions, and Processes*. Paper presented at a Symposium on Vocational Development Theory, July 1968. New York: Teachers College, Columbia Univ., 1968b. (Mimeo.)
- Thoresen, Carl E. *The Systems Approach and Counselor Education: Basic Features and Implications*. Presented to the Amer. Educ. Res. Assoc., February 1968. Stanford, Calif.: School of Education, Stanford Univ. (Mimeo.)

- Tiedeman, David et al. *Information System for Vocational Decisions: Annual Report 1966-67*. Cambridge, Mass.: Harvard Graduate School of Education, 1967.
- Tiedeman, David et al. *Information System for Vocational Decisions: Annual Report 1967-68*. Cambridge, Mass.: Harvard Graduate School of Education, 1968.
- Varenhorst, Barbara. "The Life Career Game: Practice in Decision-Making." *Simulation Games in Learning*. (Edited by Sarane S. Boocock and E. O. Schild.) Beverly Hills, Calif.: Sage, 1968.
- Wiener, Norbert. *The Human Use of Human Beings*. Boston: Houghton Mifflin, 1950.

## AUTHOR

- Cooley, William W. Address: University of Pittsburgh at Pittsburgh, Pennsylvania  
Title: Professor of Education and Computer Science Age: 38 Degrees: B.S., Lawrence College; M.A., Univ. of Minn.; Ed. D., Harvard Univ. Specialization: Computer Applications in Education; Educational Research Methodology.
- Hummel, Raymond C. Address: University of Pittsburgh at Pittsburgh, Pennsylvania  
Title: Associate Professor of Education Age: 47 Degrees: B. A., C.C.N.Y.; M.A., Columbia Univ.; Ph.D., Columbia Univ. Specialization: Counseling; Vocational Development; Educational Change.

## 10: RELEVANCE AND RESEARCH IN COUNSELING

Carl E. Thoresen\*  
*Stanford University*

In his book *Psychotherapy East and West*, Alan Watts (1961) described Samsara, the Eastern concept of life lived in a vicious cycle of endlessly repetitive attempts to solve false and unreal problems. Much of what happens in counseling "research" qualifies as Samsara. Most published research is quite simply a waste—a waste of valuable time and resources. The simple fact is that most studies, as they are conceptualized, designed, executed and analyzed, make no difference to counseling theory and practice. Published research, of course, makes other differences to investigators (e.g., personal recognition and attention), but it seldom qualifies as "disciplined inquiry" (Cronbach and Suppes, in press). As they are typically done, research studies represent a routinized, "convenience-oriented" operation (Borgatta, 1960). Calling such work research, even if it is published, does not qualify it as significant and worthy.

A colleague recently stated, "Clothes don't make the emperor," though they may proclaim him. Citing a host of previous studies, predicting a correlation or stating an hypothesis, employing a large sample and finding something statistically significant at the .05 level too often represent but conventional drapery. What merits respect is not the clothes nor the

---

\*Dr. C. Gilbert Wrenn, Arizona State University, and Dr. Willis E. Dugan, American Personnel and Guidance Association, served as consultants to Dr. Thoresen on the preparation of this chapter.



trappings nor even the title of emperor, but the man beneath and the research he does that will make a real difference in the lives of clients.

Counseling, broadly conceived as a profession devoted to assisting individuals in changing behaviors, has conducted remarkably little research which bears directly on improved methods and results of counseling. Counselors are confronted with complex and seemingly insoluble client problems—drug abuse, alienation, race relations, sex, societal anxieties, student militancy, and communication between generations. Yet counseling researchers for the most part continue to survey, to question and to relate. By saying "more research is needed!" and by doing more of the same, research promises to increase the irrelevancy and insignificance of research to theory and practice.

The field of counseling has generally lacked a disciplined research tradition, and for understandable reasons. The service demands on counselors have often been so pressing that systematic investigation into the effectiveness and efficiency of processes and products has paradoxically been ignored. "We are too busy counseling to seriously question the consequences of what we do and how we do it." In addition, counseling, like kissing, is so intrinsically interesting and satisfying that few bother to critically examine it. The field's practitioner heritage, swamped with problems and demands, has not generally viewed "research" as an endeavor with much to offer.

### *Blurring the Paradigm*

Turbulence and turmoil have been accelerating in counseling theory, research and practice. The smug complacency and calm of which Astin (1961) could write is gone. Instead "chaos prevails" in psychotherapeutic research and practice (Colby, 1964, p. 347). The field remains "... in a mess" (Rogers, 1963, p. 9), enough of a mess for Allen (1967, p. 230) to recently comment that "... counseling becomes a process in search of a consequence, a cause earnestly seeking an effect." An increasing number of counselors are seriously questioning the conceptual adequacy and procedural efficacy of existing rationales. The confusion, while discomfiting and stressful, brings with it fresh perspectives and a lessening of the tyranny of proper procedures (what is considered right and proper). This "blurring of the paradigm" (Kuhn, 1962) is permitting the emergence of new and unorthodox notions. Furthermore, it is coercing many to ask fundamental questions about cherished personal convictions. It is facilitating a "... backing off and taking a fresh look at the basic problem. . . . How may constructive changes in behavior and personality of the troubled or deviant person be facilitated?" (Rogers, 1963, p. 12). The turmoil is beginning to stimulate "extraordinary" research (Kuhn, 1962), research free from the dogmatic grip of traditional models and methodologies.

Ways to improve counseling research have been suggested. Calls have been sounded for experimental studies, replication and follow-up, operationalism, study of the "criterion problem" and overt consideration and control of variables such as client and counselor characteristics and situational factors. An outstanding publication, *Research in Counseling* (Whiteley, 1967) contains several excellent papers by Blocher, Patterson, Krumboltz and Allen (among others) on a variety of research topics. Krumboltz (1967, p. 191), for example, offered the test of relevancy for counseling research studies: "What will counselors do differently if the results of this research come out one way rather than another?"

In this chapter the reiteration of discussions found in Whiteley (1967) and related references will be avoided. Instead, a few highlights of previous chapters will be presented followed by observations on how counseling research may become more relevant to theory and practice.

### *Current Research Needs*

One major observation of every author is this issue, directly stated or implied, is that more research per se is *not* needed. Instead, a greater variety of inquiry in terms of design and content is needed.

#### *Clarity of Goals*

Several authors cited the confusion and ambiguity about goals, purposes or objectives. This confusion ranged from school guidance, student personnel and counselor education programs to specific counseling techniques with certain types of clients. How can research in counselor education, for example, concerned with problems of selecting, training, supervising, and evaluating counselors, improve the effectiveness and efficiency of training if the desired product is not clearly stated? What should counselors be able to do as a result of training? What effects should the successfully trained counselor produce? Further studies in selection, supervision and evaluation, unless attention is given to desired outcomes, were judged to be futile.

Research in effectiveness of guidance programs remained stalled, falling back on the use of descriptive surveys or checklists of services as evaluative criteria (which clearly begs the question of effectiveness) or pleading "intangible" effects. Additionally, authors repeatedly criticized studies involving counseling procedures for the irrelevance and inadequacy of outcome measures. The authors stated that there typically was little or no established relationship and rationale among treatment procedures, counseling objectives and outcome measures.

#### *Multidimensional-Multiple Criteria Studies*

The call was emphatically sounded for multivariate studies in terms of

using more than one dependent and independent variable. Studies are needed which compare several competing counseling treatments to more than one control group. Continued use of the simple one treatment-one control design, in which subjects are sometimes matched and sometimes randomly assigned to groups, was deemed archaic in light of what is now known about experimenter effects, regression artifacts, expectancy-set factors and so-called no-treatment control groups. The authors sharply criticized the use of single, global and often unrelated treatment criterion measures such as GPA or a mean change score on some personality scale. The objectives of counseling are not to change group means on personality scales, for example, but to change the actual behaviors of clients. It was pointed out that exploratory studies, in which possible effects are often unanticipated and the treatment itself is highly complex, require a variety of criterion measures to detect any possible change. Multiple measures of change at various time intervals, not just immediately after treatment, from sources other than self-report or the report of treatment counselor, were highly recommended.

### *Integrated Research*

The urgency for integration and synthesis in research was frequently cited. Theoretical rationales remain aloof from the Hows and Whys of practice. Individual research studies make unwarranted assumptions (more often presuppositions) about "all other things being equal," failing to provide evidence that any effort was made to deal with the interaction issue, i.e., how the problem, methodology and findings fit into the larger complex whole. Continued investment in single poorly conceived, casually executed and vaguely described studies offers scant dividends for theory and practice.

In contrast, the utilization of a systems approach for conceptualizing problems, designing programs and evaluating results was viewed as a most promising strategy. Such an approach offers the possibility of integrating and synthesizing the vast array of diverse rationales, techniques and research methodologies. The crisis and resultant chaos in counseling research was viewed by several authors in terms of fragmentation, theoretical insularity and the estrangement of theoretical rationales, counseling procedures and research findings.

### *Constructive Directions*

Three ideas emerge that merit consideration here: 1) the need for disciplined inquiry, 2) the need for new research models, and 3) the need for a systems research orientation.

## *Disciplined Inquiry*

Counseling researchers, for the most part, have casually borrowed research conceptualizations, strategies and methodologies from such disciplines as psychology and sociology. Researchers have given little attention to the "philosophy of science" issues involved in how research is conducted (Maslow, 1966). Research workers have generally ignored consideration of how theoretical and operational definitions markedly influence research. The logic of research operations concerning such matters as the critical relationships among a treatment procedure, its intended consequence, and how the consequence is measured has also been ignored.

Disciplined inquiry has many characteristics: 1) conducting and reporting research so the argument can be carefully examined by others, 2) conducting a sequence of studies (not just "the" study), 3) documenting carefully with detail and precision what was done and how it was done, 4) considering instrumentation in terms of the fidelity and reliability of measurement, and 5) thinking critically about what reasonable conclusions can be drawn from the data. Three closely related points of disciplined inquiry follow.

Too many researchers operate with a sterile and ritualistic conception of research. The stuff of serious inquiry, the crucial quality, lies in the thinking and rethinking, the creating and the discovering, and the continual revising of methods to match problems (Bruner, 1966; Maslow, 1966). The monolithic tradition in counseling research has perpetuated a choking provincialism, dissuading some from ever engaging in research and fostering "academic calisthenics" by others (Krumboltz, 1967).

What seems to be lacking is a willingness to ask insolent but important questions, some of which seem sensible to the asker but for which there are perhaps no answers or even no ideas of what sort of answers would be satisfactory (Bridgman, 1959). There is much need for what Toulmin (1953) called "trained imagination," rigorous and systematic speculation, and for looking at familiar phenomena in a new way.

## *The Beginning and End*

Inquiry in the disciplined sense is typically neglected at the beginning and at the end. The middle of the spectrum—(a) the initial literature survey, (b) the statement of hypotheses, (c) the defining of operations, (d) the design, and (e) statistical procedures—usually receives all of the attention. But neglect takes place at the beginning, where the demanding, often painstaking, thinking and planning is essential. This is when the investigator thinks through the question; he asks himself and his colleagues, "What can I do at this point that will best answer the question?" or "How can I frame the problem to permit the asking of empirically answerable questions?" Platt (1964), arguing for the method of "strong inference," suggested that the beginning period focus on asking what propositions are



"subject to denial" by experimental test. "How small and elegant an experiment can you perform?" The study—and it need not be an experiment—is then conducted only *after* considerable effort has been expended in determining whether the study as conceptualized and designed will ask or answer meaningful and relevant questions.

At the beginning, the researcher must often be rigorously self-critical to avoid the seduction of a favored premise and a cherished hypothesis. For the experimental researcher it is particularly important that he plan and design his experiment using multiple hypotheses, competing treatments and rival rationales. The results the investigator hopes for have a way of turning up, unless he actively seeks in the beginning to compare more than one treatment and to test multiple, contending hypotheses (Platt, 1964; Rosenthal, 1966).

The neglected end of inquiry is data analysis. Characteristically, the finding of a significant correlation coefficient or F-ratio terminates the investigation. But as Tukey (in press) and others argue, it is precisely here that the inquiry should begin in earnest—sifting through the data, turning it upside down, seeking out the unexpected, especially "unfriendly" data in terms of a favored hypothesis or preferred rationale. Such "detective work" facilitates serendipity, the discovery of important findings not anticipated. Careful analysis and reanalysis may suggest new hypotheses and provide the basis for speculation, for looking at old phenomena in new ways. (Too often, however, the detective work is misguided because the investigator is only seeking the salvation of a statistically significant difference to report.) Of course, the formal study itself must be carefully planned and well executed if such detective work is to uncover anything.

### *The Significance Game*

Counseling research, like many fields in the behavioral sciences, is one in which measurement of relevant phenomena is imprecise, uncertain, incomplete and difficult. For this reason statistics tend to flourish, in part as a substitute for the researcher's inability to control the phenomena under investigation. Statistical procedures, however, intended to be a heuristic device, may become a substitute for critical thinking and careful decision-making. A bewildering array of statistical operations have become preemptive, camouflaging imprecision, confusion and, in some cases, nonsense (Hays, 1963; Stevens, 1968). Statistics too often have had a pernicious influence on inquiry, "clothing absurdity in fully formularized attire" (Stevens, 1968). Sometimes mathematical procedures remove the investigator from the data in their true form, causing difficulty in what supportable interpretations can be made. Hays (1963), citing the perceived panacea that some investigators prescribe for statistical maneuvers, stated candidly: "If nonsense is put into a mathematical system, nonsense is sure to come out." Matters "... reach a point where our statistical procedures are substitutes instead of aids to thought" (Bakan, 1967, p. 28). Statistics may take atten-

tion away from the data of research. The data must constitute the ultimate base for any inference or conclusion.

A critical issue for research involves decisions and inferences based on statistical calculations. Several writers (e.g., Bakan, 1967; Edwards, Lindman and Savage, 1963; Hays, 1963; Lykken, 1968; Stevens, 1968) have persuasively cautioned that a significant finding in itself permits only a very limited inference. It does not stand as a measure of significance or meaning; it is not a basis for the investigator to claim that an important empirical fact has been established. An inference about the truth of an hypothesis requires the corroboration of replication, in which several studies increase confidence in the validity of a relationship, causal and otherwise. (See Lykken, 1968, for an excellent discussion of replication.)

Very little can be inferred from a single study in which a directional hypothesis is accepted at some level of statistical significance. The consequences of placebo effects, experimenter bias, and the average shared variance of so-called "unrelated" and uncontrolled variables often strongly bias the outcome in favor on the investigator and against the null hypothesis. Seldom, for example, is a sample truly random, the correlation between two variables exactly equal to 0.00, or the difference between two groups precisely zero. The investigator has often previously gathered data suggesting the direction of differences between comparison groups. The strict null hypothesis, in fact, is often believed and known to be false from the beginning (Bakan, 1967; Edwards, Lindman and Savage, 1963). Additionally, if the null hypothesis is not rejected, usually it is because the  $N$  is too small. Simply increase  $N$  if the intent of inquiry is to reject the null hypothesis (Nunnally, 1960).

The finding of a statistically significant difference, per se, is probably the least important consideration in concluding the validity of a theoretical proposition or in establishing a useful empirical fact. Far more important in determining the value of any research are 1) the coherence and reasonableness of the theory, 2) the degree of control employed, 3) the sophistication and relevance of the measuring techniques, and 4) the scientific or practical importance of the phenomena studied.

The continued preoccupation of counseling researchers with statistical significance has a stultifying effect, and it misdirects effort that is greatly needed to improve data measurement and research control procedures.

The counseling investigator is typically not seeking significant group mean differences or group average gains compared to mean changes expected in some comparable untreated group. The counseling researcher is more typically interested in change in some individual performance. He seeks to make predictions, explore causal relationships or accurately describe individual behaviors. To what extent does a certain client show successful interpersonal skills? Can a certain counselor trainee demonstrate competent interview behaviors? Given a certain career problem, how does the

individual utilize decision-making skills? Thus, the counseling researcher's unit of interest is not the aggregate, not the central tendency of a research group; it is the magnitude of change or relationship displayed by the individual subject. Focusing exclusively on group means does not legitimately permit the investigator to infer about individual performance (Bakan, 1967).

Attention in counseling research should be focused on the attainment of a particular terminal performance by a subject (Bandura, in press). The researcher must ask if the subject reached criterion as operationally defined. Too often, confusion about what specifically constitutes a successful outcome causes the investigator to seek solace in mean changes. This is somewhat analogous to the professor who has not specified what performance constitutes an "excellent" one and, therefore, uses standings on exams to assign grades, giving the top ten per cent A's, etc. Consider, for example, two counseling treatments which, when compared by using Grade Point Average, show a statistically significant finding that after treatment subjects in one treatment demonstrated a 3.5 average compared with a 3.2 average for the other group. Such findings hardly represents information of practical or relevant significance; it tends to conceal and disguise rather than to illuminate the actual differences between treatments for individual subjects. However, if outcomes of the treatment were specified in advance as to what constitutes meaningful and worthwhile individual change, then the proportion or per cent of individual demonstrating such change could be compared between treatments.

Dickenson and Truax (1966) used as one outcome measure the relative per cent of the treated and untreated underachieving college students who changed from probationary to nonprobationary (passing) status in grades. After counseling, 71 per cent of the counseled students had passing grades compared with 46 per cent of the control group students. The information on the proportion of subjects who were successful, i.e., earned passing grades, in each treatment is far more meaningful information on mean scores for both groups significantly changing from 2.81 to 3.14.

Bandura, Blanchard and Ritter (1968) used several criteria in their examination of which of three experimental treatments reduced the fear of snakes. One major outcome measure was the relative number of Ss in each group who completely eliminated their fear of snakes. What constituted complete success was established in observable performance terms prior to the actual experiment. The live modeling and participation treatment was completely successful with 92 per cent of the subjects; the symbolic modeling and systematic desensitization treatments were completely successful only with 33 and 25 per cent of the Ss respectively. None of the control Ss showed complete success. Again, such data are far more relevant than the knowledge about mean scores on anxiety scales of 1.2, 2.9, and 3.3.

Data in counseling research can speak for themselves if the focus is on actual performance of subjects. A key requirement in gathering data on



individual performance is specification of what will constitute practical or meaningful change. What will the S do; how will he behave under what conditions; and what behavior is to be considered successful (Mager, 1962)? Of course, not all types of inquiry may be concerned with attaining some clearly specified terminal performance. However, much of the research in counseling pertains to the differential effects of various treatments for particular problems. In counseling, a major concern is how does the individual change when provided specific treatment relative to how he would change without treatment.

### *Outrageous Wordology*

A person's conception of what is real and what happens is directed considerably by the words available to him. Disciplined inquiry requires careful consideration of the meaning of terms and concepts (Brodbeck, 1963). The serious researcher must offer objective referents for his terms. He cannot operate exclusively with private symbols and meanings. The novelist or poet may deliberately cultivate ambiguity, but the researcher is obligated to describe his procedures with sufficient precision that others can replicate his studies.

However, much of what is said about counseling often becomes obscured by a mystical vocabulary that defies operational definition and genuine communication (Thoresen, in press). What are the referents (the actual behaviors of clients) of terms such as *constructive growth*, *therapeutic reinforcement*, *experimental syndrome* and *autonomous, genuine and congruently related activities*? What does it mean to say that *normal counseling* was administered? Blocher (1967) recently observed that counseling has closely resembled political ideologies and religious movements; counseling has messianic figures, a coterie of followers, true believers, and the dogmatic grip of a select number of personal concepts to provide a mantle of mysticism for its operations.

Recently, more scientific sounding concepts have been advanced, e.g., *learning theory*, *reinforcement*, *reciprocal affect*. Terms such as *systematic desensitization* and *behavior therapy* are used as if they represent a single, unitary set of operations—in fact, they describe a large, diverse and sometimes conflicting array of treatment methods and rationales.

Excessive literalism in counseling contributes to the loosest of relationships among research findings, theory and practice. The effect is pernicious because persuasive and polemic logic can take precedence over empirically-based argument. One can too easily justify uncritical loyalty to his own cherished preconceptions. The antidote is to bring theoretical concepts to an empirical confrontation (Weitzman, 1967).

The problem of confused terminology suggests the need for conceptual research, inquiry addressed to concerns such as the analysis of meanings, the detection of assumptions, and the clarification of criteria (Gowin, 1961). Counseling necessarily involves itself with the individual's total behavior,



his directly observable behaviors, and his private unobservable behaviors. Unfortunately, the inquiry line has been typically drawn here, excluding from serious consideration the individual's personal reactions. This exclusion results in part from the difficulty in describing and assessing the individual's personal reactions. Vigorous conceptual inquiry is needed because it may provide the means to clarify sufficiently personal meaning and reactions for empirical study. The value of conceptual research is illustrated by Ennis's (1967) work in developing evaluative criteria for the concept of critical thinking and Sechrest and Wallace's (1967) work with the concept of adjustment.

Counseling investigators must acknowledge that logical arguments seldom solve disputes concerning either theory or practice. Words have too often been substitutes for empirical action; they offer an enveloping warmth of abstraction. Words have a way of becoming truths, and facts become confused with interpretations. Such labels as *underachievement*, *ego strength*, *hostile*, *self-concept*, and *anxious*, intended to serve as economical descriptions of a complex of behaviors, become pseudo-explanations. For example, "Yes, he often fights with others and does poorly in school . . . (because) of his poor self-concept and his extreme anxiety." Words can frequently remove one from the data of concern—the behaviors of counselors, clients and others.

Serious inquiry demands a *data* oriented language, not a prose rich with affective connotation and "purr words" (Hayakawa, 1964). A common research language is needed, and it must be clearly referenced with observable behaviors. Such a language would facilitate the consolidation of findings, the conduct of replication studies, the revision and validation of theoretical generalizations cooperative research projects. Creation of a common data language would clarify intended meanings. Of course, the use of metaphor and high-order abstractions can continue, but such usage can be so labeled and used without muddling empirical inquiry.

### *Promising Research Models*

I share with other authors in this issue a preference for multi-variate comparative experiments in which two or more treatments are compared in a variety of criterion measures. Bandura, Blanchard and Ritter (1968); Paul (1966, 1967); and Volsky et al. (1965) are excellent examples of multivariate studies. There are, however, two other types of research which deserve the serious consideration of counseling researchers: the empirical case study and the experimental-longitudinal project.

### *Empirical Case Study*

Several years ago Skinner (1958) authored a fascinating account of his early laboratory research. He began "... simply by looking for lawful processes in behavior," not knowing what he would find (p. 362). Using one and sometimes three or four subjects, Skinner meticulously observed

the results of altering specific conditions. Skinner's writing conveys a richness and excitement about what it means to be scientific ("When you run onto something interesting, drop everything else and study it . . . science is a continuous and often disorderly and accidental process."). Skinner implored researchers to move away from large samples of subjects. He argued that the use of elaborate statistics and complex experiments may prematurely structure inquiry and bypass careful and detailed observation of phenomena. He urged researchers to reduce the sample size, to carefully control conditions, and to observe change.

Research in counseling has typically sidestepped the careful control and systematic observation of phenomena which must serve as the bedrock for hypotheses and generalizations (Samler, 1968). Descriptive studies abound, but they do not report the controlled systematic observations of individual behaviors. Such studies merely describe loosely such things as group means and correlations of test responses with other test responses or questionnaires.

A descriptive analysis of research articles reported in the *Journal of Counseling Psychology* during the last two years—part of a study in progress at Stanford of published research in several counseling journals by Carl E. Thoresen and David Shaw—found that over 75 per cent of the research articles published were descriptive-correlational reports. None reported controlled observation and change of individual behaviors. The remaining 25 per cent of research articles (approximately one-third of all articles were not research reports) reported experimental studies. Of these experimental studies, about 70 per cent used a single criterion, 50 per cent used one experimental group, 20 per cent used randomization of subjects to treatments (over 60 per cent "arbitrarily" assigned Ss to treatments), and 75 per cent used a "posttest only" design. This latter finding is instructive because it indicates that no individual pretreatment assessment of criterion behaviors occurred. Hence, inferences about magnitude of change as a result of treatment for individuals remain sharply curtailed, regardless of what criteria were used. (See Campbell and Stanley, 1963; Solomon and Lessac, 1968.)

The empirical case study is not meant to be synonymous with the traditional clinical case study which is characterized by uncontrolled observations without measurement (Bolgar, 1965; Paul, 1967). The empirical case study can be viewed as combining the careful measurement and observation of the highly controlled laboratory procedures with the real clients, counselors and problems of field research. The free operant conditioning studies described by Bijou and Baer (1966) share much with the empirical case study: 1) attention to and control of the physical and social environment, 2) concern with the single individual, 3) focus on observable changes in the individual over time, and 4) use of response frequencies or response magnitudes as measures of change.

What commends the empirical case study at this point in counseling research is its focus on the direct observation and recording of behavior in life as well as quasi or simulated settings. Such data can yield a rich data base from which hypotheses can be generated and tested in subsequent experiments. A most promising characteristic of this method centers on its clinical application. The practicing counselor can utilize the empirical case method with individual clients. This method permits the counselor to engage in a type of highly relevant inquiry, since empirical data are provided on client behaviors as those behaviors change (or fail to change) during treatment. In addition to providing valuable information for the counselor on the relationship between treatment procedures and changes in client behavior, the empirical case study serves as an important innovative tool. Data may be gathered, for example, about new techniques and their influence on particular clients. Such data may provide the basis for generating testable hypotheses in subsequent experimental studies.

A related variation of the empirical case study is the time-series design quasi-experiment (Campbell and Stanley, 1963). The interrupted time-series design, for example, allows the counseling researcher to assess the effects of a particular treatment on a subject over time (Campbell, 1963) (see reference to Briskin and Gardner, 1968, Chapter 4, this issue). A variation of this is the "A-B-A own-control design," sometimes called intra-subject replication study (Bandura, in press), in which some target behavior is repeatedly altered from one direction to another, contingent upon the application of a counseling procedure. Another variation entails the alternate uses of a change technique using pre- and posttreatment measures each time the treatment is administered.

I am currently planning to use the empirical case method with college students who suffer extreme social anxiety. Various self-report, observer and physiological measures, obtained with telemetry equipment, will be gathered from individual subjects in a variety of life settings. The ongoing effect of treatment procedures will be observed and recorded.

The empirical case method cannot answer questions about the comparative effects of two or more treatments and about the value of a particular treatment for different types of clients. However, it represents a valuable method for directly observing and gathering relevant data and examining the effects of a particular intervention or sequence of interventions over time. Further, it will often prove to be an effective antidote to premature and superficial experimentation which is designed and executed before the availability of relevant data on individual performance.

#### *Experimental-Longitudinal Project*

This research model, recently advanced by Bijou (1968) (see also Bijou, Peterson and Ault, 1968), offers the best of two worlds: the precision and control of the cross-sectional experiment and the descriptive richness and follow-through of the longitudinal study. Examples of such work are



the fraternal-twin study of Dennis (1941) and Hilgard's (1932) study of academic training on achievement of preschool children. Much experimentation, no matter how well designed and rigorously executed, leaves unanswered questions about generalization effects, stability of change, and interaction effects between specific changes in client behavior over time with changing environmental conditions. In contrast, longitudinal studies describe but cannot establish functional relationships. Such data, however, can be suggestive of causal processes and can provide critical data on changes over time. The marriage of cross-sectional experimental and longitudinal strategies will permit the combining of several research techniques, such as questionnaires, interviews, formal experiments and time-series studies. The question asked or the problem explored will determine which technique or combination of techniques should be used.

Although cross-sectional studies (e.g., Krumboltz and Thoresen, 1964; Thoresen and Krumboltz, 1968) have produced evidence that various social modeling and reinforcement procedures significantly increased information-seeking behaviors, many problems remain unsolved. For example, some students engaged in considerably more information-seeking behaviors than did others. Longitudinal data about the base rate performance of students on this variable and other behaviors could provide information suggestive of differential counseling treatments. Data are needed on how the acquisition of career-related information subsequently influences a host of other behaviors that have some bearing on personal life, school and career. How does a particular treatment such as group social modeling interact with particular student characteristics over time to influence other changes? Subjects should be carefully assessed on a variety of measures before, during and at several times following the treatment. Exclusive reliance on cross-sectional experimentation will not answer some of these important questions.

A rich source of data can be provided by the use of an experimental-longitudinal strategy in which all subjects are carefully followed over time, some subjects are readministered the same treatment while other subjects are exposed to other experiences. New subjects can be introduced in the project at several points and administered treatments while original subjects undergo various combinations of assessment and treatment procedures over time. This paradigm, of course, requires a team of investigators.

It is time to end the parochialism in counseling research that employs only one strategy. Just as there is no one best counseling method for every client, so is there no one best research method for every problem. The empirical case study and experimental-longitudinal research seem to be particularly useful ways for investigating many problems.

### *A Systems Perspective*

Teilhard de Chardin (1959) has written in *The Phenomenon of Man*



that the more man penetrates and explores matter, the more he is confounded by its intricate complexities and the interdependence of its parts. Much of what confronts the counseling researcher is interdependent and complex in ways still unknown and not understood. Counseling researchers have often been forced to make reductionistic assumptions (Koch, 1964; Maslow, 1966) about their subjects, treatments and outcomes. Many counselors have, in turn, dismissed most research as naive and nonsensical because their clinical experiences have revealed the complexities of client behaviors. In the last two decades, a systems perspective has emerged from a variety of sources; a few of these sources are general systems theory, engineering, theoretical biology and political science. The notion of a "system" is centuries old, but the systems approach is new in that it centers on 1) studying a system as an integrated, interdependent entity rather than as a conglomeration of parts or elements and 2) reorganizing and changing components of a system to produce objectives or goals stated in observable performance terms.

A system can be defined as a structure which functions as a whole by virtue of the interdependence of its parts. The concepts of input, processes and output are common to a systems approach. A systems approach, in general, can be viewed as a design for action, a conceptual framework for establishing, analyzing and synthesizing alternative ways of achieving clearly stated systems objectives. An eye is constantly directed toward maximizing benefits and minimizing costs. (See Cooley and Hummel, Chapter 9, this issue; Thoresen, in press; Von Bertalanffy, 1967; Kuhn, 1963; and Buckley, 1968 for a more thorough explication of the systems approach). In general, a systems perspective represents a way of thinking and viewing complex problems. The emphasis is clearly on the "big picture," i.e., on identifying all of the relevant variables and factors, especially how and to what effect they mutually interact.

A serious limitation of contemporary research in counseling is its fragmentation and unrelatedness: theory here, practice there, and research "somewhere else." Contemporary theory and research have most noticeably failed to provide a rationale for the *prevention* of problems and the *development* of optimal behavior. Research strategies at best are directed toward providing data for decisions and inferences about remediation of existing client problems. Research on what can be termed "developmental prevention" is lacking. A broadening conceptualization of theory and research, such as that suggested by Blocher (1967), is vital if counseling is to increase appreciably in effectiveness in serving a heterogeneous clientele.

A systems orientation offers a conceptual scheme for the conduct of more integrated and synthesized research. Such a framework may permit the bringing together of the conglomerate of counseling procedures, rationales and related developments. For example, a vast array of "machines" such as portable audio and video recorders, physiological telemetry

instruments, and digital and analog computers are available, waiting to be integrated into effective man-machine combinations in counseling, guidance and counselor preparation programs. The pioneer work of Tiedeman (1968) and his associates at Harvard in vocational decision-making illustrates the coordinated use of man-machine combinations. What has already been suggested here—disciplined inquiry, need for a data language and operational terms, multivariate studies, empirical case studies and experimental-longitudinal research—fits well within a systems orientation.

Arguments about the value of "process versus outcome" and "basic versus applied" research are blunted within a systems framework. The desired output, the product, determines what research activities are most appropriate in terms of sequence, content and methodology. A systems framework by definition incorporates so-called process studies with outcomes and merges "basic" (findings of interest primarily to scientific colleagues) and "applied" (findings of greater interest to practitioners). Given a particular set of objectives, all levels and types of inquiry may be appropriate.

There are many exciting, challenging developments and ideas of relevance to counseling and guidance that demand the systematic attention of counseling researchers. Leonard (1968) in *Education and Ecstasy* outlined a wealth of provocative notions about new environments for learning which have many implications for the practice of counseling, guidance programs and counselor education. Rogers (1968) advanced "encounter groups" as the most significant social innovation of the century. Yet this development and other group methods as suggested by Schutz (1967) and Gunther (1968) have not received disciplined attention. Cohen (1968) and Goldiamond (1968) provided new vistas in "behavioral architecture," the creation of integrated physical and psychological environments which effectively and efficiently promote learning. Krumboltz and Thoresen (1968) are initiating a new system of counselor education drawing on systems theory and the behavioral sciences. A systems research approach promises the perspective and breadth necessary to conduct serious inquiry in these areas.

Within the field there are some exciting beginnings that suggest that the time is right for the objective-focused, mission-oriented, comprehensive framework provided by a systems perspective: Blocher's (1967) conceptions of life stages, life spaces and life styles, suggesting a complex interactive orientation; the "micro-counseling" simulation studies of Ivey et al. (1968) in preparing counselors; the interpersonal process recall and stimulus film techniques of Kagan and Schauble (in press); and the behavioral-analytic model for assessing personal problem-solving competence of college students (Goldfried and D'Zurilla, in press). Generally, the value of a systems perspective is underscored by the increasing acceptance of the question "What treatment, by whom, is most effective for this individual with that specific problem?"

Research, like counseling itself, can no longer remain one-method, one-discipline and one-person oriented if it is to deal with contemporary problems demanding serious inquiry. Marcuse (1964, p. 256) wrote that "The struggle for solution has outgrown the traditional forms." A systems research perspective promises to provide the flexible and comprehensive forms needed.

### Conclusion

Counseling research should evolve from the problems and concerns of counselors and their clients. There must be an "intimate and genuine relationship" in counseling between the what and how of research, the why of theory and the wherewithal of counseling practices.

Relevance in counseling research will be enhanced by:

- (1) initiating inquiry with focus on the problems and concerns of contemporary man;
- (2) exploring how the hypotheses and data of the behavioral sciences and technical fields, such as experimental social psychology and computer technology, suggest concepts and methods for new counseling techniques and rationales;
- (3) focusing more attention on specific problem-relevant changes in the behaviors of individual subjects rather than on less related mean score changes on psychometric devices;
- (4) utilizing a wider array of research methodologies such as the experimental case study, experimental-longitudinal projects and systems research perspective; and
- (5) acknowledging the obstacles of counseling terminology and moving toward a common empirical data language.

Counseling, a profession dedicated to preventing and ameliorating human problems, must demand that its research make a difference in the lives of clients. The action in research needs to be where the client is—his thoughts, his feelings, his actions—in sum, his behaviors.

### Bibliography

- Allen, Thomas W. "Task Group Report One." *Research in Counseling*. (Edited by John M. Whiteley.) Columbus, Ohio: Charles E. Merrill Publishing Co., 1967. Chapter 7, pp. 219-237.
- Astin, Alexander. "The Functional Autonomy of Psychotherapy." *American Psychologist* 16: 75-78; 1961.
- Bakan, David. *On Method: Toward a Reconstruction of Psychological Investigation*. San Francisco: Jossey-Bass, 1967.
- Bandura, Albert. *Principles of Behavior Modification*. New York: Holt, Rinehart and Winston, in press.
- Bandura, Albert; Blanchard, Edward D.; and Ritter, Brunhilde J. *The Relative Efficacy of Desensitization and Modeling Approaches for Inducing Behavioral, Affective and Attitudinal Changes*. Stanford, Calif.: Stanford Univ., 1968. (Mimeo.)



- Bijou, Sidney W. "Ages, Stages and the Naturalization of Human Development." *American Psychologist* 23: 419-423; 1968.
- Bijou, Sidney W. and Baer, Donald M. "Operant Methods in Child Behavior and Development." *Operant Behavior—Area of Research and Application*. (Edited by Werner K. Honig.) New York: Appleton-Century-Crofts, 1966. Chapter 17, pp. 718-789.
- Bijou, Sidney W.; Peterson, Robert F.; and Ault, Marion H. "A Method to Integrate Descriptive and Experimental Field Studies at the Level of Data and Empirical Concepts." *Journal of Applied Behavior Analysis* 1: 175-191; 1968.
- Blocher, Donald. "What Can Counseling Offer Clients? Implications for Selection." *Research in Counseling*. (Edited by John M. Whiteley.) Columbus, Ohio: Charles E. Merrill, 1967. Chapter 1, pp. 5-29.
- Bolgar, Hedda. "The Case Method." *Handbook of Clinical Psychology*. (Edited by Benjamin B. Wolman.) New York: McGraw Hill, 1965. Chapter 2, pp. 28-39.
- Borgatta, Edgar F. "A Commentary on Small Group Research." *Decisions, Values and Groups*. (Edited by Dorothy Willner.) New York: Pergamon Press, 1960, pp. 179-184.
- Bridgman, Percy W. *The Way Things Are*. Cambridge: Harvard University Press, 1959.
- Brodbeck, May. "Logic and Scientific Method in Research on Teaching." *Handbook of Research on Teaching*. (Edited by N. L. Gage.) Chicago: Rand McNally, 1963. pp. 44-93.
- Bruner, Jerome S. "Some Elements of Discovery." *Learning by Discovery*. (Edited by Lee S. Shulman and Evan R. Keislar.) Chicago: Rand McNally, 1966. Chapter 7, pp. 101-113.
- Buckley, Walter (editor). *Modern Systems Research for the Behavioral Scientist*. Chicago: Aldine, 1968.
- Campbell, Donald T. "From Description to Experimentation: Interpreting Trends as Quasi-Experiments." *Problems in Measuring Change*. (Edited by Chester W. Harris.) Madison: University of Wisconsin Press, 1963. Chapter 12, pp. 212-242.
- Campbell, Donald T. and Stanley, Julian C. "Experimental and Quasi-Experimental Designs for Research in Teaching." *Handbook of Research on Teaching*. (Edited by N. L. Gage.) Chicago: Rand McNally, 1963. Pp. 171-246.
- Cohen, Harold L. "Behavioral Architecture: New Thinking for New Universities." *Training Professionals in Procedures for the Establishment of Educational Environments. A Report on the CASE Training Institute*. (Compiled by H. L. Cohen, et al.) Silver Spring, Md.: Educational Facility Press, Institute for Behavioral Research, 1968. Pp. F3 1-11.
- Colby, Kenneth Mark. "Psychotherapeutic Processes." *Annual Review of Psychology* 15: 347-370; 1964.
- Cronbach, Lee J. and Suppes, Patrick (editors). *Disciplined Inquiry in Education*. New York: Macmillan Co., in press.
- Dennis, Wayne. "Infant Development Under Conditions of Restricted Practice and of Minimum Social Stimulation." *Genetic Psychology Monographs* 23: 143-189; 1941.
- Dickenson, Walter A. and Truax, Charles B. "Group Counseling with College Under-achievers." *Personnel and Guidance Journal* 45: 243-247; 1966.
- Dunnette, Marvin D. "Fads, Fashions and Folderol in Psychology." *American Psychologist* 21: 343-352; 1966.
- Edwards, Ward; Lindman, Harold; and Savage, Leonard J. "Bayesian Statistical Inference for Psychological Research." *Psychological Review* 70: 193-242; 1963.
- Ennis, Robert H. "Readiness to Master a Principle." *Psychological Concepts in Education*. (Edited by B. Paul Komisar and C. B. J. Macmillan.) Chicago: Rand McNally and Co., 1967. Pp. 51-59.
- Goldfried, Marvin R. and D'Zurilla, Thomas J. "A Behavioral-Analytic Model for Assessing Competence." *Current Topics in Clinical and Community Psychology*. (Edited by Charles D. Spielberger.) New York: Academic Press, in press.
- Goldiamond, Israel. "Architectural Space and Media in the Planning of Learning Processes." *Training Professionals in Procedures for the Establishment of Educational Environments. A Report on the CASE Training Institute*. (Compiled by H. L. Cohen et al.) Silver Spring, Md.: Educational Facility Press, Institute for Behavioral Research, 1968. Pp. F2 1-21.



- Gowin, D. Bob. "Prospects for Philosophic Inquiry in Guidance." *The School Review* 69: 191-205; 1961.
- Gunther, Bernard. *Sense Awareness*. New York: Macmillan Co., 1968.
- Hayakawa, S. I. *Language in Thought and Action*. Revised edition. New York: Harcourt, Brace and World, 1964.
- Hays, William L. *Statistics for Psychologists*. New York: Holt, Rinehart and Winston, 1963.
- Hilgard, Josephine R. "Learning and Maturation in Preschool Children." *Journal of Genetic Psychology* 41: 40-53; 1932.
- Ivey, Allen E. et al. "Microcounseling and Attending Behavior: An Approach to Practicum Counselor Training." *Journal of Counseling Psychology, Monograph Supplement* 15: No. 5, Part 2: 1-12; 1968.
- Kagan, Norman R. and Schauble, Paul. "Affect Simulation in Interpersonal Process Recall." *Journal of Counseling Psychology*, in press.
- Koch, Sigmund. "Psychology and Emerging Conceptions of Knowledge as Unitary." *Behavioralism and Phenomenology*. (Edited by T. W. Wann.) Chicago: University of Chicago Press, 1964.
- Krumboltz, John D. "Future Directions for Counseling Research." *Research in Counseling*. (Edited by John M. Whiteley.) Columbus, Ohio: Charles E. Merrill Publishing Co., 1967. Chapter 6, pp. 184-203.
- Krumboltz, John D. and Thoresen, Carl E. "The Effect of Behavioral Counseling in Group and Individual Settings on Information-Seeking Behavior." *Journal of Counseling Psychology* 11: 324-333; 1964.
- Krumboltz, John D. and Thoresen, Carl E. *Behavioral Systems Training Program for Counselors*. Stanford, Calif.: School of Education, Stanford Univ., May 1968. (Mimeo.)
- Kuhn, Alfred. *The Study of Society: A Unified Approach*. Homewood, Ill.: Dorsey Press, 1963.
- Kuhn, Thomas S. *Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.
- Leonard, George. *Education and Ecstasy*. New York: Delacorte Press, 1968.
- Lykken, David T. "Statistical Significance in Psychological Research." *Psychological Bulletin* 70: 151-159; 1968.
- Marcuse, Herbert. *One-Dimensional Man*. Boston: Beacon Press, 1964.
- Mager, Robert F. *Preparing Instructional Objectives*. Palo Alto, Calif.: Fearon, 1962.
- Maslow, Abraham H. *The Psychology of Science*. New York: Harper and Row, 1966.
- Nunnally, Jum C. "The Place of Statistics in Psychology." *Educational and Psychological Measurement* 20: 641-650; 1960.
- Paul, Gordon L. *Insight Versus Desensitization in Psychotherapy*. Stanford, Calif.: Stanford University Press, 1966.
- Paul, Gordon L. "The Strategy of Outcome Research in Psychotherapy." *Journal of Consulting Psychology* 31: 109-118; 1967.
- Platt, John. "Strong Inference." *Science* 146: 347-353; 1964.
- Rogers, Carl R. "Psychotherapy Today or Where Do We Go From Here?" *American Journal of Psychotherapy* 17: 5-16; 1963.
- Rogers, Carl R. *The Increasing Involvement of the Psychologist in Social Problems: Some Comments Positive and Negative*. Paper presented to the Amer. Psychol. Assoc., September 1968. La Jolla, Calif.: Western Behavioral Science Institute.
- Rosenthal, Robert. *Experimenter Effects in Behavioral Research*. New York: Appleton-Century-Crofts, 1966.
- Samler, Joseph. "The Counselor in Our Time." *Rehabilitation Counseling Bulletin* (Special Issue) 11: 198-287; 1968.
- Schutz, William C. *Joy—Expanding Human Awareness*. New York: Grove Press, 1967.
- Sechrest, Lee and Wallace, John. *Psychology and Human Problems*. Columbus, Ohio: Charles E. Merrill, 1967.
- Skinner, B. F. "A Case History in Scientific Method." *Psychology: A Study of a Science*. (Edited by S. Koch.) New York: McGraw-Hill, 1958.
- Solomon, Richard L. and Lessac, Michael S. "A Control Group Design for Experimental Studies of Developmental Processes." *Psychological Bulletin* 70: 145-150; 1968.
- Stevens, S. S. "Measurement, Statistics and the Schemapiric View." *Science* 161: 849-856; 1968.

- Teilhard de Chardin, Pierre. *The Phenomenon of Man*. New York: Harper, 1959.
- Tiedeman, David V. "Economic, Educational and Personal Implications of Implementing Computerized Guidance Information Systems." *Information Systems for Vocational Decisions*. Project Report No. 13, Cambridge: Harvard School of Education, March 1968.
- Thoresen, Carl E. "The Counselor as an Applied Behavioral Scientist." *Personnel and Guidance Journal*, in press.
- Thoresen, Carl E. "The Systems Approach and Counselor Education: Basic Features and Implications." *Counselor Education and Supervision*, in press.
- Thoresen, Carl E. and Krumboltz, John D. "Similarity of Models and Clients in Behavioral Counseling: Two Experimental Studies." *Journal of Counseling Psychology* 15: 393-401; 1968.
- Toulmin, Stephen. *The Philosophy of Science: An Introduction*. New York: Harper and Row, 1953.
- Truax, Charles B. "Discussion." *Research in Counseling*. (Edited by John M. Whiteley.) Columbus, Ohio: Charles E. Merrill, 1967. Pp. 208-214.
- Tukey, John W. "Analyzing Data: Sanctification or Detective Work." *American Psychologist*, in press.
- Volsky, Theodore, Jr. et al. *The Outcomes of Counseling and Psychotherapy: Theory and Research*. Minneapolis: Univ. of Minn. Press, 1965.
- Von Bertalanffy, Ludwig. *Robots, Men, and Minds*. New York: Braziller, 1967.
- Watts, Alan. *Psychotherapy, East and West*. New York: Pantheon Books, 1961.
- Weitzman, Bernard. "Behavior Therapy and Psychotherapy." *Psychological Review* 74: 300-317; 1967.
- Whiteley, John M. "Refocus of Counseling Research." *Research in Counseling*. (Edited by John M. Whiteley.) Columbus, Ohio: Charles E. Merrill, 1967. Chapter 11, pp. 261-278.

#### AUTHOR

Thoresen, Carl E. Address: Stanford University at Stanford, California Title: Assistant Professor Age: 35 Degrees: A.B., Univ. of Calif., Berkeley; A.M., Stanford; Ph.D., Stanford Specialization: Experimental Research in Counseling Techniques; Small Group Counseling; Systems Approaches in Counseling and Counselor Education.

for counseling—  
for administrative use—  
**and also**  
**for research—**

---

measures of individual  
and group characteristics

DAVIS READING TEST

ACADEMIC PROMISE TESTS

SURVEY OF STUDY HABITS & ATTITUDES

WECHSLER INTELLIGENCE SCALES

DIFFERENTIAL APTITUDE TESTS

AND MANY OTHERS!  
WRITE FOR COMPLETE TEST CATALOG

**NEW**  
OMNIBUS PERSONALITY INVENTORY

Free to AERA members — two new research reports — the Junior College Validity Study and the Junior College Normative Data Study — with validity data and norms for the College Qualification Tests, the Davis Reading Test, and the Survey of Study Habits and Attitudes.

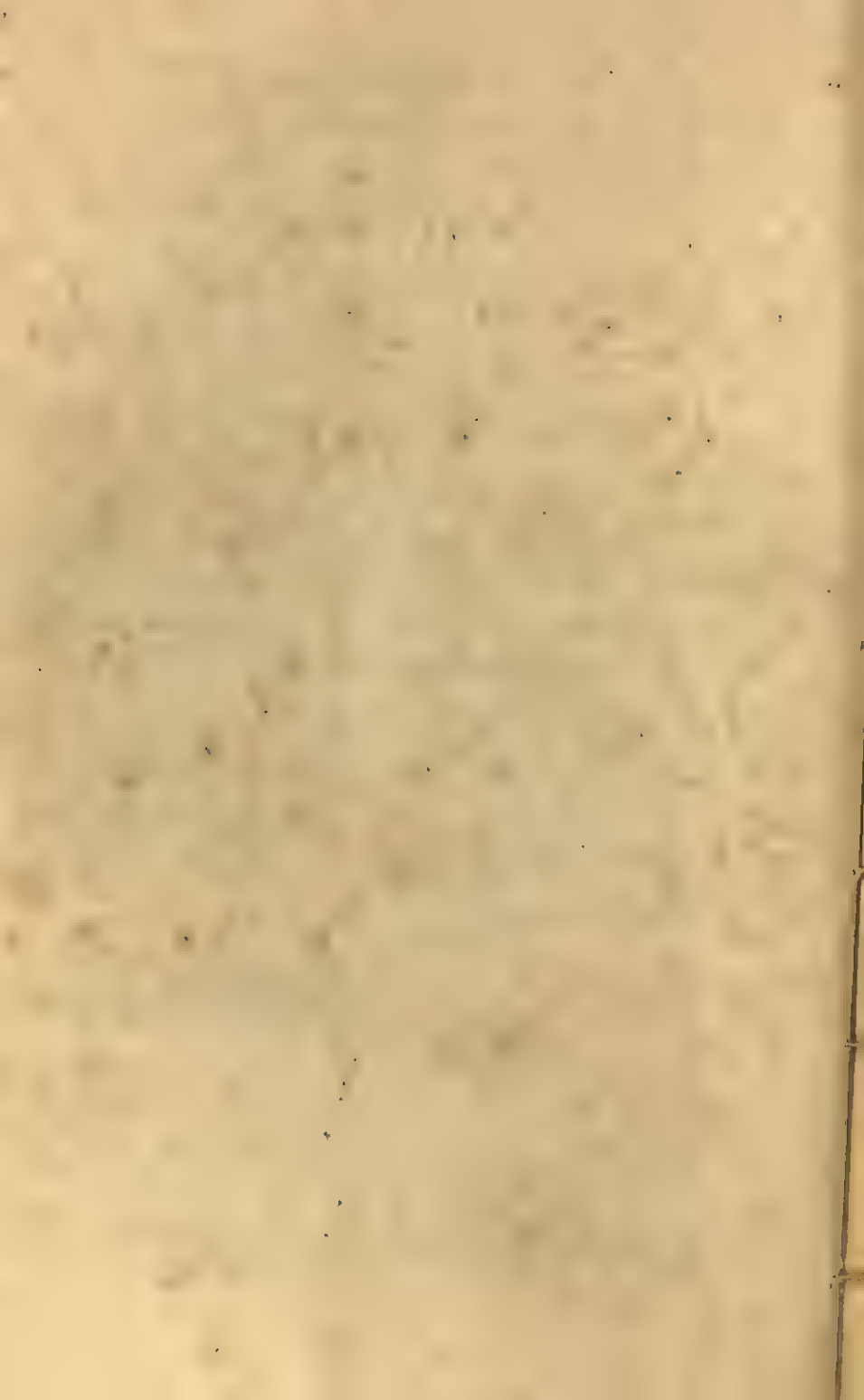
**Ψ The Psychological Corporation**  
318 East 45th Street, New York, New York 10017

## FOREWORD

In our opinion, curriculum research represents one of the most potentially influential areas of educational research activity. The high probability that curriculum research will one day markedly affect educational practice makes us particularly grateful to our distinguished associates who undertook, as chapter authors and chapter advisers, the preparation of this issue of the *Review*. In addition to the four initial chapters which were based on an examination of curriculum research, the latter two chapters should be of particular interest to the reader. A special invited chapter dealing with methodological issues in curriculum research was prepared by Richard E. Schutz. The final chapter by John I. Goodlad represents the comprehensive considerations of an eminent curriculum scholar. We regret sincerely that a chapter concerning research on curriculum organization was not prepared by the designated author in time for inclusion in this issue. Those who took time to react to the four initial chapters warrant our special thanks. Finally, we are most grateful to Gene V Glass, the Editor, and to his associates for the excellent suggestions proffered regarding both the form and substance of the chapters.

W. James Popham  
Robert L. Baker  
Issue Editors





# 1: HISTORY OF CURRICULUM THOUGHT AND PRACTICE

ARNO A. BELLACK\*

*Teachers College, Columbia University*

Curriculum innovators of the past decade have attempted to solve the difficult problems of curriculum planning and development with scant attention to the historical dimensions of these problems. The Committee on the Role of Education in American History (1965) contended that "imperfect knowledge" of educational history had "affected adversely the planning of curricula . . . in the continuing crisis of American education." Goodlad (1966, p. 91), in his critique of the contemporary curriculum reform movement, deplored the fact that "A substantial number of the new crop of reformers have approached the persistent, recurring problems of curriculum construction in the naive belief that no one had looked at them before."

This ahistorical stance seems to be characteristic not only of the current crop of curriculum reformers, most of whom are university professors of academic disciplines, but also of educationists who claim curriculum building as their field of professional specialization. Kliebard (1968, p. 69) maintained that in the curriculum field "issues seem to arise *ex nihilo*; each generation is left to discover anew the persistent and perplexing problems that characterize the field." He contended that the inability of curriculum specialists to see their field in perspective resulted in a "tendency to repeat the rallying cries and slogans that had their origins in a different intellectual climate and a different social milieu as if they had an immediacy that they no longer possess" (Kliebard, 1968, p. 69). He concluded that if curriculum is to prosper as a field of study, critical examination of inherited ways of thinking about curriculum problems is essential. He urged curriculum theorists and practitioners to engage in dialogue not only among themselves, but also with their professional forebears.

There is evidence of dawning interest on the part of at least some historians and curriculum specialists in participating in such a dialogue. For example, Cremin (1966), an historian, viewed the current curriculum reform movement as "essentially continuous with the efforts of the early progressives." He contended that the ultimate aim of both reform efforts

\*Dr. Mauritz Johnson, State University of New York at Albany, served as the consultant to Dr. Bellack on the preparation of this chapter.

is the same: "to humanize knowledge so that it can be popularized." And Goodlad (1968), a curriculum specialist, examined the "excesses" both of the progressive reformers of the 1930's and 1940's and of the discipline-centered innovators of the 1950's and 1960's; he made proposals for the future which would integrate the achievements of the progressive and discipline-centered eras.

Interest in the contributions that history might make to the study of curriculum is beginning to develop not only in the United States, but in England as well. Charlton (1968), a British historian, discussed the relevance of the historian's methods (basic concepts and modes of investigation) to curriculum theorizing; he explored the question of the extent to which the content of history can be drawn on in dealing with problems of curriculum planning, not to provide answers or solutions to problems, but "to make us aware of the possibility of change, of the complexity of change, and of the carryover of the past into our present situation and future aspirations" (p. 77).

### *Formative Years of the Curriculum Field*

For centuries educational philosophers and theorists have given serious attention to basic curriculum questions (what to teach, how to teach, and how to improve teaching), but the self-conscious identification of certain educational theorists and practitioners as curriculum specialists is a development of twentieth-century American education. The emergence of curriculum as a field of professional work and study was explored by Seguel (1966), Caswell (1966), and Kliebard (1968a).

Seguel examined certain aspects of developing thought about curriculum problems from the 1890's to the late 1930's that contributed to the establishment of curriculum as a specialized field of study. The views of seven influential men, viz., Charles McMurry, Frank McMurry, John Dewey, Franklin Bobbitt, W. W. Charters, Harold Rugg, and Hollis Caswell, were examined to cast light on the outlines of emerging thought that eventually led to the awareness by the educational profession of the nature and importance of curriculum making. The writings of these key figures were analyzed to discover how each one perceived the educational situation, and why he became interested and involved in the study of curriculum; what proposals he made about the curriculum; and how relevant to educational practice his proposals proved to be. Seguel concluded that the period under study could be characterized by four persistent interests on the part of the emerging group of curriculum specialists: 1) the nature of knowledge, 2) the nature of the knowing process, 3) the professional status of the new specialty of curriculum making, and 4) procedures for introducing new curriculum insights into educational practice on a broad scale.

Caswell (1966) presented an account of the "organized curriculum movement" as it took form during the 1920's and 1930's. Since that time, the persistence of certain educational problems that "are studied systematically only by general curriculum workers" has, according to Caswell's interpretation, made the curriculum a field of established professional importance. He identified three continuing, central concerns of curriculum specialists: 1) assuring sound sequence or continuity in the curriculum, 2) establishing consistent relationships between general goals of education and specific objectives that guide teaching, and 3) designing curricula that provide a reasonable balance of emphasis among the various areas of study.

Kliebard (1968a) also saw the 1920's as the decade in which curriculum emerged as a field of professional activity. He explored the intellectual climate and prevailing modes of thought of the age in which Bobbitt achieved prominence as the foremost practitioner in the curriculum field. The predominant social doctrine espoused was "social efficiency," which in curriculum terms "held up all school subjects, indeed all school activity, against the criterion of social utility" (Kliebard, p. 75). The key idea was that subjects taught in the schools were to be judged by criteria external to the subjects themselves. In addition to the criterion of social utility applied to school studies, Kliebard pointed out that the social efficiency movement incorporated two closely related dichotomies that were to have a profound effect on twentieth-century curriculum development: 1) the dichotomy of school subject—the academic and the practical, and 2) the dichotomy of school population—college preparatory and non-college preparatory.

Critical studies that trace the development of the curriculum field beyond its formative years are much needed. Of special importance would be investigations of the 1940's, when the war caused severe dislocations in the schools, and the 1950's, in which a new breed of reformers led a vigorous reform and wrested leadership from the curriculum specialists who traced their professional lineage to Bobbitt and Charters. Such inquiries should focus attention not only on the development of curriculum proposals and doctrines championed by individuals and organizations, but also on the impact of these proposals on actual educational practices.

### *Curriculum Theorists*

In an essay on the historiography of American education, Cremin (1965, p. 79) contended that "one of the great lacks to date has been a sufficient number of detailed, critical studies of individual educators." To a limited degree, this need is being met by a number of studies of educators who influenced the development of curriculum thought and practice.

Wirth (1966) critically appraised one period in John Dewey's career



when he had direct and continuing responsibility for the development of an educational program. This was the decade 1894-1904 when Dewey was chairman of the combined Departments of Philosophy, Psychology, and Education and founder-director of the University Laboratory School at the University of Chicago. Wirth examined the influence of Dewey's general philosophical theory on the curriculum and methodology developed in the Laboratory School. Dewey's "Plan of Organization of the University Primary School," prepared shortly before the Laboratory School opened, appeared in print for the first time in an appendix of Wirth's study.

A collection of lectures given by John Dewey on the Philosophy of Education at the University of Chicago in 1899 was edited by Archambault (1966). These lectures constituted, in the editor's words, "the pilot study for the *Democracy and Education* of 1916." They were concerned not only with criticism of educational practice and rebuttals of the views of Herbartianism, the reigning doctrine of the time, but also with Dewey's own emerging views on curriculum and teaching methods. Archambault detected a "striking consistency" between the educational theory set forth by Dewey in these lectures and the clarification of his mature position in *Experience and Education* in 1938.

Drost (1967a) presented a professional biography of David Snedden, spokesman in the early decades of the twentieth century for "the most thoroughgoing form of social efficiency." In many respects, Snedden's curriculum proposals anticipated the life adjustment ideology of the mid-1940's and the early 1950's. Drost argued that although Snedden's popularity began to decline in the late 1920's, many of the views he advanced were adopted by others and have continued to the present time to be important alternatives in educational planning, particularly in certain contemporary proposals for disadvantaged students.

Harold Rugg's efforts to reform the teaching of the social studies were investigated by Winters (1967). He dealt primarily with the writing and publication of the unified social science textbook series which Rugg chose as the vehicle to carry his ideas into American classrooms. Rugg's curriculum was based on "great central concepts" related to the crucial problems of social life that had been identified by the "frontier thinkers" in the social sciences.

Eisner (1967a) analyzed Bobbitt's efforts in the 1920's to demonstrate how scientific principles might be applied to the practical problem of curriculum construction. He described the educational context in which Bobbitt worked—an environment in which the ideas of social efficiency, scientific management, experimentalist theory, and psychological measurement were widely proclaimed. Eisner examined the consequences that followed from Bobbitt's approach to curriculum building and explored the reasons why Bobbitt's work had little impact on the progressives, the essentialists, and rational humanists. Eisner identified certain "family

resemblances" between the views of Bobbitt and certain contemporary curriculum theorists including Ralph Tyler, Benjamin Bloom, Virgil Her-  
rick and John Goodlad.

A cogent reappraisal of James B. Conant's role in recent educational history was presented by McClellan (1968). He traced the development of Conant's views on public education, including his views on curriculum issues, from the 1930's to the present. Using the tools of contemporary philosophical analysis, McClellan examined the arguments that Conant presented for his curriculum proposals and concluded that "Conant has not provided and could not provide rationally compelling arguments for educational policy."

### *National Curriculum Commissions*

Since the latter part of the nineteenth century, curriculum making by national committees has been a distinctive feature of American education. Historical investigations of these commissions—the educational and cultural milieu in which they worked, the nature of their policy proposals, and their influence on classroom teaching—are essential to understanding the development of curriculum thought and practice.

The reports of the Committee of Ten in 1893 and of the Commission on the Reorganization of Secondary Education in 1918 were landmarks in the development of the high-school curriculum.Sizer (1964) presented an appraisal of the report of the Committee of Ten—its setting, its composition, its distribution, and its effects. Sizer viewed the report of the Committee as "the first such effective document in the history of American education." The Committee, whose specific function was to bring about reforms in the instructional program, "standardized the main-line subjects in the secondary school curriculum." Sizer contended that after initial success in influencing the development of high-school education, the report became obsolete in two decades, due largely to a rapidly expanding school population and far-reaching social and cultural changes in the nation. Krug (1964) reappraised both the role of the Committee of Ten and the influence of the Commission on the reorganization of Secondary Education in the shaping of the American high school. Krug's volume is indispensable reading for an understanding of how the American high school came to be the institution it is.

There was hope in some quarters that the report of the Committee of Fifteen in 1895 which dealt with elementary education would do for the elementary schools what the Committee of Ten had done for the high schools. This turned out to be an idle hope, as Button (1965) indicated in his assessment of the Committee of Fifteen. Button explored the reasons why the doctrines of the Committee of Fifteen were rejected for other views more acceptable to influential reformers of the time. Drost (1967b) presented an account of the "great debate" occasioned by the presentation

of the report of the Committee of Fifteen. He expressed the view that the controversy at the heart of the debate was "the issue of content selected for specific purpose versus the integrity of the subject fields" and suggested that this issue has turned out to be a continuing dilemma for curriculum makers.

National curriculum commissions exercise no less influence in the contemporary reform movement than they did in earlier periods. Under varied auspices, including national professional organizations of the academic disciplines, national commissions have been at work during the past decade designing curricula to revitalize the teaching of many of the subjects in the school program. A number of these commissions have been at work long enough to permit critical examination of their activities in an historical perspective.

Wooton (1965) presented an account of activities of the School Mathematics Study Group from its initiation in 1958. He analyzed the various projects of SMSG and explored the reasons why the crucial decision was made to concentrate initially on the writing of sample textbooks. DeMott (1964) investigated the skirmishes during the period of 1961-1964 in the "math wars" between the opponents and proponents of the new mathematics.

Glass (1964) traced the curriculum improvement activities of the Biological Sciences Curriculum Study during the years 1959-1964; he concluded that during its first five years the BSCS program incorporated two distinctive features that previous efforts to improve the science program had failed to stress: first, large numbers of research scientists collaborated with high school teachers in replacing antiquated content with contemporary scientific knowledge, and second, the new curriculum stressed the understanding of the nature of scientific inquiry rather than the acquisition of scientific information.

Critical historical studies of the activities and policy proposals of the numerous national curriculum commissions that have been at work during the past decade are definitely needed. Many of these commissions have introduced into the mainstream of educational thought proposals that have profoundly influenced the course of curriculum development in the schools, and there is no understanding the present state of affairs apart from their contributions.

### *Curriculum Problems and Issues*

To say that the contemporary curriculum problems have historical roots is to be guilty of a commonplace. But given the pervasive ahistorical posture of the curriculum field, it is a truism that curriculum specialists would do well to keep in mind. In this section a limited number of studies that help place certain current issues in historical perspective are reviewed.



The present era of curriculum development is marked by widespread interest in programmed instruction. Dale (1967) reviewed certain past developments in curriculum and teaching that provide an historical context for some of the persistent problems that have to be dealt with in introducing innovations like programmed instruction into the school program. He described innovations championed by reformers in the past that included certain principles and characteristics now combined in programmed instruction: activity analysis and specification of behavioral objectives; criterion tests of terminal behavior; feedback on the results of learning efforts and instructional designs; individualized instruction; and educational engineering. Dale's analysis revealed the continuing influence on contemporary developments in programmed instruction of early leaders in curriculum building like Bobbitt and Charters, who stressed activity analysis and detailed specification of objectives in behavioral terms.

Questions about the role of objectives in instructional planning are persistent concerns for curriculum specialists. Eisner (1967b) and Kliebard (1968b) traced the interest of curriculum makers in stating objectives in behavioral terms to the work of Bobbitt in the 1920's, and both argued that the doctrine of behavioral objectives has serious limitations.

Eisner examined Bobbitt's viewpoint on the central importance of specific behavioral objectives in instructional planning and showed that the influence of Bobbitt's views extends to the present in the work of certain curriculum theorists such as Ralph Tyler, Benjamin Bloom and David Krathwohl. Eisner contended that the theory of behavioral objectives has serious limitations because it has "not sufficiently emphasized the extent to which the prediction of educational outcomes cannot be made with accuracy"; and it has not adequately taken into account "the ways in which the subject matter affects precision in stating educational objectives." Eisner's essay was followed by critical comments by three specialists in the fields of curriculum and testing (Ebel, 1967; Hastings, 1967; Payne, 1967); this intriguing interchange of views closed with a response by Eisner.

Kliebard (1968b) explored the social and educational context in which the "ideology" of behavioral objectives emerged in the 1920's with the work of Bobbitt and traced the influence of this viewpoint through the intervening years to contemporary curriculum theories of Ralph Tyler and Will French. Kliebard (p. 246) suggested several points at which he considered the notion of behavioral objectives to be most vulnerable, the most serious of which was that "from a moral point of view, the emphasis on behavioral goals . . . still borders on brainwashing or at least indoctrination rather than education."

The views of Eisner and Kliebard regarding the value of defining objectives in behavioral terms differ markedly from those expressed by Dale (1967). Although the issues at stake in the dispute obviously cannot



be settled by historical analysis, such analysis does provide useful background and perspective for consideration of the questions under debate.

During the 1960's, the empirical study of the teaching process has become a major interest of educational researchers. Broudy and Palmer (1965) and Lee (1966) provided an historical setting for current concerns with teaching. Broudy and Palmer concentrated on teaching method as it was exhibited in the work of a selected number of noted teachers from ancient times to the recent past, including Socrates, Abelard, Comenius, Pestalozzi, Froebel, Herbart, and Kilpatrick. The authors were primarily concerned with conveying something of the teaching style of these "exemplars" and how their teaching reflected or exemplified the pedagogical problems of their age. The basic thesis of their volume was that "it is the success routes of an era that dictate the dominant patterns of schooling and the styles of teaching."

Lee analyzed the ways in which the role of the American teacher and common perceptions of that role have changed in recent years. He focused upon teaching and teachers in the period since the outbreak of World War II and attempted to compare conceptions of the late 1930's and the 1940's with those of the mid-1960's. He concluded that "we have been moving since the war toward a more modest and manageable conception of the teacher's function."

Contemporary theories of curriculum and teaching method have been influenced by many diverse conceptions from the past, including "the project method" which enjoyed great vogue during the heyday of progressivism. Bleeke (1968) reviewed the development of the project method from its early beginnings in agricultural programs to its later identification with the activity curriculum. Major attention was given to key figures such as W. H. Kilpatrick and J. F. Hasic.

Today it is generally agreed that if programs of curriculum development are to influence classroom behavior, provision must be made for the active participation of teachers. Peltier (1967) raised the question, "How did the classroom teacher first acquire responsibility for planning the curriculum?" He chose the Denver Public Schools for an historical case study of teacher participation in curriculum revision. The period studied was the 1920's, when the Denver approach to curriculum making stressed not only the involvement of teachers, but also the use of curriculum specialists and the establishment of a permanent curriculum department.

A persistent need in any field of study is precision in meaning of its basic terminology. In curriculum literature few terms have been so inadequately defined as the term "experience curriculum." Phillips (1965) attempted to clarify the meanings of the term by examining its historical development since the 1930's as revealed in the writings of prominent exponents of the experience curriculum.

## Afterword

The historical investigations reviewed in this brief chapter are evidence of interest on the part of at least some contemporary historians and curriculum specialists in studying the history of curriculum thought and practice. Those who pursue historical studies in this area would do well to keep certain guiding ideas in mind. First, historical inquiry should not be viewed as a search in the past for solutions to present-day instructional problems. As was suggested earlier in this chapter, the purpose should be to help make us aware of the possibility and complexity of curriculum change and conscious of the carryover of past doctrines and practices into the present situation. Second, the historian would do well to avoid what has been called "the sin of evangelism" (Cremin, 1965)—attempting to inspire teachers with professional zeal, rather than helping them understand what actually happened in the development of curricula in the schools. Third, it should be borne in mind that the history of curriculum thought and practice cannot be separated from the general history of American education, which, in turn, cannot be divorced from the broader stream of cultural and intellectual history.

## Bibliography

- Archambault, R. D. (editor). *Lectures in the Philosophy of Education: 1899 by John Dewey*. New York: Random House, 1966.
- Bleeke, Milton. *The Project: From a Device for Teaching to a Principle of Curriculum*. Doctor's Thesis. Madison: Univ. of Wis., 1968.
- Broudy, Harry S. and Palmer, John R. *Exemplars of Teaching Method*. Chicago: Rand McNally, 1965.
- Button, H. W. Committee of Fifteen. *History of Education Quarterly* 5: 253-63, 1965.
- Charlton, Kenneth. *The Contributions of History to the Study of the Curriculum. Changing the Curriculum*. (Edited by J. F. Kerr.) London: Univ. of London Press, 1968. Pp. 63-78.
- Committee on the Role of Education in American History. *Education and American History*. New York: Fund for the Advancement of Education, 1965.
- Caswell, H. L. *Emergence of the Curriculum as a Field of Professional Work and Study. Precedents and Promises in the Curriculum Field*. (Edited by H. F. Robinson.) New York: Teachers College Press, 1966. Pp. 1-11.
- Cremin, Lawrence A. *The Genius of American Education*. New York: Random House, 1966.
- Cremin, Lawrence A. *The Wonderful World of Ellwood Patterson Cubberly*. New York: Bureau of Publications, Teachers College, Columbia Univ., 1965.
- Dale, Edgar. *Historical Setting of Programed Instruction. Programed Instruction. Sixty-sixth Yearbook, Part II, National Society for the Study of Education*. Chicago: Univ. of Chicago Press, 1967. Chapter 2, pp. 28-54.
- DeMott, Benjamin. *The Math Wars. New Curricula*. (Edited by R. W. Heath.) New York: Harper and Row, 1964. Pp. 54-67.
- Drost, Walter H. *David Snedden and Education for Social Efficiency*. Madison: Univ. of Wis. Press, 1967a.
- Drost, Walter H. *The Immortal Day in Cleveland—The Report of the Committee of Fifteen. Educational Theory* 17: 178-91; 1967b.
- Ebel, Robert L. *Some Comments. School Review* 75: 261-66, 1967.
- Eisner, Elliot W. *Educational Objectives: Help or Hindrance? School Review* 75: 250-60, 277-82; 1967b.

- Eisner, Elliot W. Franklin Bobbitt and the "Science" of Curriculum Making. *School Review* 75: 29-47; 1967a.
- Glass, Bentley. *Renascent Biology: A Report on the AIBS Biological Sciences Curriculum Study. New Curricula.* (Edited by R. W. Heath.) New York: Harper and Row, 1964. Pp. 94-119.
- Goodlad, John I. *The Changing School Curriculum.* New York: Fund for the Advancement of Education, 1966.
- Goodlad, John I. Curriculum: a Janus Look. *The Record* 70: 95-107; 1968.
- Hastings, J. Thomas. Some Comments. *School Review* 75: 267-71; 1967.
- Kliebard, Herbert. *The Curriculum Field in Retrospect. Technology and the Curriculum.* (Edited by P. W. F. Witt.) New York: Teachers College Press, 1968a. Pp. 69-84.
- Kliebard, Herbert. Curricular Objectives and Evaluation: a Reassessment. *High School Journal* 51: 241-47; 1968b.
- Krug, Edward A. *The Shaping of the American High School.* New York: Harper and Row, 1964.
- Lee, Gordon C. The Changing Role of the Teacher. *The Changing American School. Sixty-fifth Yearbook, Part II, National Society for the Study of Education.* Chicago: Univ. of Chicago Press, 1966. Chapter 1, pp. 9-31.
- McClellan, James. *Toward an Effective Critique of American Education.* Philadelphia: J. B. Lippincott Co., 1968. Chapter 2, James Bryant Conant: A Man Made System and Vice-Versa, pp. 59-127.
- Payne, Arlene. Some Comments. *School Review* 75: 272-77, 1967.
- Peltier, Gary L. Teacher Participation in Curriculum Revision: an Historical Case Study. *History of Education Quarterly* 7: 209-19; 1967.
- Phillips, Richard C. The Historical Development of the Term, Experience Curriculum. *History of Education Quarterly* 5: 121-30; 1965.
- Seguel, Mary Louise. *The Curriculum Field, Its Formative Years.* New York: Teachers College Press, 1966.
- Sizer, Theodore R. *Secondary Schools at the Turn of the Century.* New Haven: Yale Univ. Press, 1964.
- Winters, Elmer A. Man and His Changing Society: The Textbooks of Harold Rugg. *History of Education Quarterly* 7: 493-514; 1967.
- Wirth, Arthur G. *John Dewey as Educator.* New York: John Wiley and Sons, 1966.
- Wooten, William. *SMSG, The Making of a Curriculum.* New Haven: Yale Univ. Press, 1965.

# AUTHOR

- ARNO A. BELLACK Address: Teachers College, Columbia University at New York, New York Title: Professor of Education Age: 35 Degrees: B.S., Univ. of Wis.; M.A., Univ. of Wis.; Ed.D., Teachers College, Columbia Univ. Specialization: Research on Teaching and Curriculum.

## 2: FORCES INFLUENCING CURRICULUM

JOHN D. McNEIL\*

*University of California, Los Angeles*

Changes are occurring both in theories for constructing curriculum and in what is actually taught in the schools. Knowledge of the powerful forces influencing these changes is limited. An indication of the diverse approaches used deriving such knowledge and the results attained from these efforts will be found in this chapter. Most of the studies cited are valuable for their curricular implications and because the findings reported are of general interest.

Forces include overlapping categories of social and economic life, e.g., class and race tensions, legislation, private foundations and technology. Groups and individuals—United Mexican-American Students, American Legion, Conant, Bruner, Zacharias—perceived as being influential may also be forces. Similarly, powerful ideas and products are forces whether or not the particular individuals responsible for their development are generally recognized. Cases in point are the recently drawn distinctions between summative and formative evaluation (Scriven, 1967), the idea of item sampling (Lord, 1962), schemes for task analysis (Gagné, 1967), and the Coleman Report (1966). Many of these forces are influencing the curriculum of the schools. Fewer forces are modifying the way that curriculum specialists study the process for planning educational programs and instructional sequences, i.e., curriculum as a field of knowledge regarding how the curriculum should be made. There is little evidence that principles of curriculum and instruction (knowledge in the process of curriculum making) are significant in shaping the curricula of the schools. Neither are there many studies relating specific instructional practices to a given societal force or condition. In the absence of studies demonstrating causality, I have cited, in the section that immediately follows, research showing strong associations between selected forces and the development of policy, which in turn is presumed to influence the curriculum. Also included in this section are reports and speculations which purport to identify shaping trends. In a final section the reader will find mention of the forces which are operating in the field of curriculum itself (curriculum as a field of inquiry). I have reported those ideas and practices which promise to

---

\*Dr. Harold Baker Dunkel, University of Chicago, served as the consultant to Dr. McNeil on the preparation of this chapter.



influence the research strategies of those who claim to have knowledge of more effective ways to plan curriculum.

## *Association of Forces With Educational Policy And Curriculum Change*

### *Some Bases for Decisions Affecting the Curriculum*

Gideonse (1968) made an analytical study of research to be used at policy-making levels by commissions, state legislatures or the U. S. Congress. His study is important for several reasons. He differentiated the kind of research needed for curriculum decision making by controlling agencies (e.g., school board members) as opposed to research necessary for decisions made by those operating at the instructional level (e.g., teachers). Gideonse described the kinds of investigation that should be conducted to make research a more powerful tool in the development and implementation of broad educational policy, and he identified those present studies that are likely to be used in making policy decisions which affect the curriculum. Among the studies seen by Gideonse as being the most valid for decision making are the following:

(1) *The Equal Educational Opportunity Survey* (Coleman, 1966). As a result of this study there will be greater efforts to advance social integration of pupils in schools and to improve teacher preparation. These projected emphases derive from Coleman's finding that characteristics of the teaching staff and social characteristics of the student body correlated more highly with student achievement than with other factors such as facilities and curriculum when home family background was statistically controlled.

(2) *Project TALENT* (Flanagan, 1962). One of the conclusions drawn from data of Project TALENT is that school size, average size of classes, age of building and suburban locations are unlikely to be important causes of excellence of school output. It might be inferred from this conclusion that less money should be spent for such purposes as new schoolhouse construction. Perhaps funds would have more important instructional effect if they were spent on something like supervision of instruction. It is possible, however, that pressures other than those for pupil achievement (e.g., the need for stimulating the construction industry or the desire to increase real estate values) will dominate educational decisions.

(3) *The Plowden Report* (1967). One of the many interesting findings in this report from England is the discovery that parental attitudes have greater correlation with pupil achievement than material home circumstances or variations in school and classroom organization, instructional materials and particular teaching practices. This finding can be used to buttress the case for helping schools overcome instructional opposition

now operating outside the school environment. School efforts can, for instance, be reinforced by arranging for parental participation in classroom instruction.

(4) *Stability and Change in Human Characteristics* (Bloom, 1964). Bloom reviewed and analyzed approximately one thousand longitudinal studies of early learning. A central conclusion of the Bloom study is that changes in the development of human characteristics become more difficult with increasing age. Perhaps Bloom's finding is having an impact upon support for pre-school programs and is responsible for much of the recent heavy emphasis upon developing curriculum and instruction for very young children.

### *Some Manifestations of Forces*

In "The Changing Curriculum," Shaw (1966) focused attention on a number of forces believed to be determining the direction of innovations in curriculum. Shaw cited such stimulants to change as advances in technology, concerns for education of the culturally disadvantaged, governmental programs and special interest groups. Newer instances of these forces can be added:

(1) The Scholarly Pressure Group. The Research and Policy Committee of the Committee for Economic Development (1968) is calling for more investment in research, development and innovation; it emphasizes four "imperatives": (a) better organization for innovation and change, (b) more emphasis on basic and applied educational research, (c) employment by school systems of the results of cost benefit and cost effectiveness analysis to decide how to allocate the resources available to education, and (d) establishment of a national commission on research, innovation and evaluation in education with the objective of bringing about more effective methods of instruction.

(2) Popular Social Protests. Franklin (1968) reported that in recent months virtually every organization working for the advancement of the Negro American, whether militant, moderate, or in-between, has demanded a more adequate treatment of Negro Americans in books and courses dealing with the history of the United States. Hechinger (1968) listed organizations and their educational demands associated with the nation's mood of angry activism.

(3) Private Foundations and Government. The common citizen may think that he has no difficulty seeing a possible relationship between the interests of private foundations and particular curricular emphases. He may believe, for instance, that the Ford Foundation's programs are straightforward in their emphasis on the disadvantaged as shown by awards to schools which plan to prepare high school dropouts for another try at school, inner-city job placement programs, television shows on the racial crisis, recruitment and training of teachers for inner-city schools. However,

the targets of these funded projects may be more than those labeled as "disadvantaged." As a case in point, the school organization itself and, in turn, all children, are Fantini's (1967) real target. Fantini, Program Associate of the Ford Foundation, made it clear that it is institutional change that should be sought rather than either the financial strengthening of the existing educational system or mere treatment of the obvious casualties of this system, namely, the poor and the culturally deprived. Fantini wants federal efforts (Title I of the Elementary and Secondary Act 1965, Project Head Start, Upward Bound, and other established O.E.O. projects) to serve as vehicles for institutional change. Although federally-funded programs would be aimed nominally at helping the disadvantaged, these programs could be designed to bring profound benefits to all. To obtain funds, the school would have to show how spending the money would result in a changed organizational structure to allow for experimentation, innovation and collaboration between various institutions. Schools could maximize their chances for getting federal monies by showing that programs designed for special populations of learners also carry provisions for wider use of community resources in curriculum planning, broader considerations in framing policy for student admission, teacher collaboration in curriculum planning, hiring of nonprofessional aides to teachers, and letting private corporations apply their "know-how" in the school.

(4) Technology. C. C. Anderson (1968), drawing heavily from Galbraith (1967), looked at two developments he thought would influence what goes on in schools—technology and the contemporary industrial state. He depicted unpleasant consequences from the merging of these influences and suggested a new remedy.

Although technology has brought many mechanical instruments into the classroom, Black (1968) wrote that it is an error to conclude that the textbook has lost its primacy as a teaching tool. He argued that the textbook is still the basis for every curriculum and that it determines what will be taught and when and, to a large extent, how. Black also described how textbooks come into being, the pressures exerted by publishers and the influences brought to bear on publishers, the effects of censorship, and the power of a few states to force changes in textbooks. A recent study in which the influence of technology was minimized is Oettinger's (1968) "The Myths of Educational Technology." Oettinger questioned whether as a result of such inventions as systems analysis and the computer we can now make "important additions to current practice." He expressed doubts that there is sufficient necessary curriculum knowledge for redesigning a school program. Oettinger stated that "We even lack a useful set of empirically validated principles of instruction" and pointed out the problems that must be solved before technology can have an impact on progress in the curriculum.



(5) **Social and Psychological Trends.** A collection of papers from a conference treating influences in curriculum change is available from ASCD (1968). Authors of these papers have tried to identify the social and psychological forces which could form the curriculum of the future. Foshay (1968) saw five trends shaping the curriculum of the future: (a) a re-emerging concern with general education; (b) new developments in economical school practices (e.g., "within ten years, new ways of thinking about who should study what and for how long will be available"); (c) class and race tensions; (d) influence of education in other lands; and (e) effect of legislation (action programs are being supported at the expense of research). Bricker (1968) reported psychological studies of specific techniques of reinforcement to get children and others to behave in desired ways. Although this research shows the existence of effective reinforcers which could be used to make the instructional process more effective, the author makes no claim that the potential of these forces (techniques) will indeed be released in the classroom.

#### *Process of Influence and Evidence of Effect*

Few researchers have traced the process in which societal forces or conditions are transmitted to instructional practices and material. Fewer investigators have attempted to demonstrate the effects of a particular force upon either intervening changes such as school curricular offerings or changes in the learner himself. From the vantage point of his historical study,Sizer (1964) concluded that schools do change, but they seem to change only when both the gap between schools and the society is extreme and the demand for formal education is growing. Those who have completed studies of lesser scope—lesser in both range of effect and effect over time—report faculty and administrative personnel arrangements associated with an increase in the school's inventory of materials and programs. Torkelson and Driscoll (1968) summarized many such current studies. One of these studies is by Godfrey (1965), who found no standard or typical channel for influencing the use of audio-visual media in the school. School boards and parents were not influential; teacher requests, demonstrations and observations of operation programs were more effective. Because the Godfrey study and those like it are descriptive and rely on correlational data to imply relationships, they lack information about how a school board or parents can change their procedures and acquire greater influence in the classroom. Sources for designing experimental studies on this latter question can be found in literature about the change process in the schools: the work of Brickell (1961) and Doll (1964); and special publications by ASCD (1966) and by The Center for the Advanced Study of Educational Administration (1965).

The survey continues to be popular for assessing the potency or impotency of forces. Cawelti (1967) reported the number of schools that accepted or rejected 27 particular innovations in a national survey of



7,237 accredited high schools. According to Cawelti's survey, the typical high school has adopted an average of six of these innovative practices within recent years. The most popular innovations were the use of the CHEM study curriculum (38.7 per cent), SMSG mathematics (36.4 per cent) and PSSC physics materials (43.2 per cent). With the exception of language laboratories, there appeared to be limited use of educational technology. The most frequently reported innovations in organization were team teaching (41 per cent), employment of teacher aides (29.1 per cent) and college credit courses (28 per cent). A high abandonment rate for innovations such as certain new science (3 per cent) and math (6 per cent) curriculums, instructional television (4 per cent), programed instruction (5 per cent) and team teaching (4.3 per cent) requires further study. In his most recent survey, Conant (1967) found that only 10.8 per cent of 2,000 schools fulfilled all five of the minimum academic requirements which he set earlier (Conant, 1959). Conant concluded that in the ten years since he made his first study considerable progress had been made in the offering of mathematics, science and foreign language. No evidence has been gathered as to the impact of the academic inventory authored by Conant in the earlier decade. Neither is the influence of Bruner's *The Process of Education* (1960) known. Did it cause change or was it used to sanction changes that would have occurred anyway? Bruner's book, with its statement of faith that children can learn any subject in an intellectually honest form at any age, has been one of the most quoted sources. Bruner reported (1968):

The book touched off a debate on the problem of revising Italian education to cope with the changing industrial society, and it has been used for clubbing Marxists and classicists alike. In the Soviet Union, one group of social critics has used the book's emphasis on discovery and intuition to castigate the dogmatism of remaining Stalinists who wish to set the dogma of socialism on the line in the classroom. . . . In Japan, the social critics praise the book for indicating that school subjects that are technical and mathematical need not be without a proper intellectual structure and cultural grace. In Israel, a land surrounded by a ring of hostile nations, the book has been greeted as an invitation to avoid mediocrity in the preparation of new immigrants—a mediocrity that social critics fear will bring Israel to a state of dangerous vulnerability in her present isolated position. In the United States—and perhaps this is the only country affluent enough to harbor such thoughts—the principal social criticism has been a concern for the maintenance of spontaneity of the child. It has been a sobering experience to realize in what degree a book of this sort must perforce serve social and political ends and can never remain a technical book alone.

In addition to describing and analyzing curriculum materials prepared by a large number of projects, Goodlad et al. (1966b) presented a summary of shaping forces—social factors operating in the years immediately after World War II, the accumulation of knowledge, federal leadership, private foundations—forces lying largely outside the state and local systems charged legally with responsibility for determining what to teach. This work is cited, however, chiefly because the authors illuminated some of the problems and issues that arise from attempts at curricular reform and placed these problems within a broader context.

### *Comment*

What is the significance of the studies cited thus far? What can be said about their trustworthiness as research? The studies above were not selected because they are exemplars of research, but because they represent what is now available to show and explain curricular influence. When asked about the influence of the famous Eight-Year Study (the first of the national programs with which he has been associated), Tyler (1967b) cautioned against attributing change to any specific influence unless the relationship can be clearly demonstrated. Few such relationships have been shown. There is a lack of experimental investigations in which predicted changes in learners occur as a result of manipulation of a force. Most of the studies cited fail to show a linkage or a differentiation between forces affecting society as a whole, education as a field, and curriculum as it is actually found in the classroom. Some are journalistic or prophetic; their authors identify trends as shaping the curriculum not on the basis of data but because the trends reflect the direction the author wants or does not want to see the schools go. Researchers do not make clear why, when pursuing relations to curriculum, they select technology, court decisions, the newspaper, the College Entrance Examination Board, or any other of the many alternative factors available. Neither is there evidence that the list of factors chosen for investigation is comprehensive. There is no apparent theory guiding these particular inquiries; no structure for coordinating individual efforts. On the positive side, the authors have offered interpretation beyond that found in popular literature by providing data that forces heralded as powerful may not be as influential as supposed, and, conversely, that some forces have not been recognized for the full effect they may be capable of making.

### *Forces Within the Field of Curriculum*

The study of curriculum is more than describing particular courses or programs of studies; it should provide knowledge of better ways to answer perennial questions of what and how to teach. Three publications in particular have been instrumental in summarizing and guiding the theory-building, research and activity found in this field in which learning

more about how to conduct the process of curriculum making is a central goal (Tyler, 1950; University of Chicago, 1950; and Smith, Stanley and Shores, 1950). Increasing attention to the development of curriculum as disciplined inquiry with its own problems, methods, and sources of data can be noted in each of the three latest issues of the *Review of Educational Research* on "Curriculum Planning and Development" (1966, 1963, 1960). Goodlad and Richter (1966a) extended the curriculum ideas of Ralph Tyler in an effort to systematically identify and relate the critical problems of curriculum to each other. In this work, they pulled from the past and present the questions which seem to constitute the substance of curriculum theory, and they cast them in what is hoped will be more productive orders. Careful to regard curriculum—intended learnings—as only part of the educational culture, the authors concentrated on what is involved in selecting, justifying, and arranging these learnings. They also made distinctions among the kinds of decision and the data believed to be necessary for producing a rational curriculum. To the extent that the system proposed is used to motivate and guide both theoretical inquiry and research, its potential as a force within the field will be fulfilled. Talmage (1967) found evidence that theoretical models for conceptualizing a curriculum system can be effectively used in guiding curriculum research design.

### *Formulating Instructional Objectives*

#### I. Subject Matter as a Data Source.

The most recent review of curriculum planning and development (AERA, 1966) pointed out that both theorists and practitioners were trying to define the "structure" of various "disciplines" and to construct curriculum materials in accordance with these definitions. Determination of both ends and means of the curriculum was dependent largely on what was said by scholars in a given field and by philosophers of science—those who specialized in analyzing and describing the fields of knowledge themselves (Ford and Pugno, 1964; Phi Delta Kappa, 1964; The Edward A. Uhrig Foundation, 1961). There were concurrent reactions against the new subject matter approach to curriculum making. Fears were expressed that objectives formulated chiefly in response to considerations of available organized intellectual resources (i.e., knowledge as the irreducible element—not learner or social concerns) would have undesirable consequences. One potential consequence was said to be an increased fragmentation of subject matter and the strengthening of compartmentalized courses at the expense of comprehensive and integrated knowledge. In response, some curriculum theorists formulated broad conceptual schemes to expose learners to a comprehensive range of disciplines and suggested how knowledge from these fields could be brought to bear on social and personal problems (King and Brownell, 1966; Broudy et al., 1964; Phenix, 1964).



## II. Society as a Data Source.

Studies of contemporary life have been a second source of data for deriving objectives. In industry, for example, objectives for training programs are based upon the critical analysis of activities carried on by those who are already fulfilling the future role expected of trainees. Mastery of the skills actually employed on the job becomes the objective of the training program. Similarly, in the past, school people analyzed the citizens' activities in health, family, economics, leisure time, and civic and vocational life to be sure that school objectives were relevant to demands of life outside of school (Bobbitt, 1918; Rugg, 1926). Now once again within the field of curriculum, society as a data source for deriving objectives is a controversial topic and an area of research. Kliebard (1968) cautioned against drawing objectives from the limitless realm of human activity instead of looking for objectives within a field of study and the classroom process. He doubted that classroom instruction can have a cause-and-effect relationship upon one's behavior in situations remote from the classroom. Husén (1965) acted upon an opposing assumption. Husén empirically selected curriculum objectives that are relevant to the needs of modern life and work in Sweden. He (a) expressed educational objectives in terms of skills and knowledge which the adult will need as a worker and as a citizen—not exclusively the knowledge metaphysically anchored in an academic field and (b) established standards reached in school and determined how long these standards are maintained in later life.

There is an issue as to whether instructional objectives of all school offerings should be empirically validated by their effects upon the learner's subsequent performance or whether instructional objectives should be philosophically justified as ends but not experimentally manipulated as means. Recently, McGuire and Kersh (in press) analyzed a series of statistical studies and found that the frequently claimed negative relationship between driver education courses in high schools and the number of serious accidents was apparent, not real. In response to this finding, there were some arguments that prevention of accidents should not be the objective of driver education any more than acquiring expertness at games should be the objective of physical education. This kind of response dramatizes the pressing issue of means and ends in curriculum.

As might be expected, much of the research activity in formulating objectives on the basis of social conditions is being pursued in vocational education. Phipps and Evans (1968) referred to a number of studies on objectives which serve more than one particular work situation. Most of these studies were guided by a "cluster concept" in which it is assumed that it is better to acquire the skills and knowledge for a family or cluster of occupations than those only appropriate for a particular job which may not exist in the future.



In contrast to reacting to immediate social demands such as protest against inequality of opportunity, some scholars are more concerned with the future and are basing instructional objectives on the implications drawn from the kinds of future envisioned. A fruitful source of ideas and methods for making more reflective decisions regarding education in the future is found in Hirsch (1967).

### III. The Learner as a Data Source.

In his article describing what he taught in a summer school, Borton (1967) distinguished between the clichés about students' *interests* and *concerns*, their basic psychic and social drives. A student might be interested in cars because he feels powerless. How to help a learner explore his understanding of power might become a springboard for curriculum development. Borton attempted to develop and carry out a curriculum of concerns. It included, among other things, having the children identify the disparity between what they thought about in school, what they were concerned about in their own lives, and the way they acted. Perhaps this project will help to stimulate curriculum workers to define better what is meant by deriving objectives from studies of the learner.

How can data about learners' interests, concerns, and academic and social deficiencies be best used in formulating objectives? When is study of the learner necessary to gain control over him and when is it undertaken to help him fulfill his own desires? It is customary to present the learner with a problem, note his response to it and if the response is judged inadequate, to say that the learner "has a need," a gap to be closed. How do objectives that are formulated on the basis of data showing gaps in a learner's behavior in some situations or class of situations differ from those formulated on the basis of assumptions about the basic drives of students? The attainment of a long-range objective imposed socially may require teaching a hierarchy of prerequisite objectives to those lacking the prerequisites. In such cases, the prerequisite objective should not be labeled as one formulated by the learners. Similarly, taking account of learner characteristics to plan and organize instruction more effectively toward an imposed objective cannot be called self-determining. Presently, most curriculum projects aimed at a particular kind of learner, e.g., potential dropouts (Mathews et al., 1966) and "unmotivated" rural youth (Phipps and Fuller, 1965), are more of the gap-closing variety than the self-initiating.

#### *Stating Objectives*

Central among tools for classifying objectives and indicating their order of complexity have been the *Taxonomy of Educational Objectives, Handbook I: Cognitive Domain* (Bloom, 1956) and *Handbook II: Affective Domain* (Krathwohl, 1964). Simpson (1967) developed a taxonomy for the psychomotor domain. The influence of Mager's *Preparing Objectives*

for *Programed Instruction* (1961), a programed text characterized by its demand for behavioral objectives and close association of the objectives with criterion measures, is seen in newer, locally constructed courses of study and in lesson plans (Smith, 1968). It should be remembered, however, that what is a force within the field of curriculum, e.g., belief in the importance of stating objectives behaviorally, is not necessarily a force in the implementation of curricula in classrooms. Ammons (1964), for example, produced evidence that (a) some school systems do not have objectives to guide their educational programs, (b) school systems do not follow recommended curriculum process to develop their instructional objectives, and (c) teachers base their programs on what they customarily have done rather than on the system's educational objectives. An indication that the inattention to guidelines can be altered is found in Popham and Baker's study (1965) showing that curriculum principles can be taught to teachers to affect their uncued responses.

Gagné (1965) presented an excellent critical discussion of the problem of stating objectives and the importance of the problem in the design of effective instruction. He reported research upon the effect of specifying instructional objectives and stated inadequacies found in the *Taxonomy of Educational Objectives, Handbook I* (Bloom, 1956). Edling (1968) reviewed a number of studies showing the influence of media upon the clarification of educational objectives.

In his article "New Dimensions in Curriculum Development," Ralph Tyler (1966) stated that the level of generality appropriate for an objective is perhaps the most puzzling question about objectives that curriculum workers currently face. Other indications that the subject of specificity of objectives is a current issue in curriculum can be found in the writings of Eisner (1967) and Paulson (1967).

Researchers have not shown a workable method by which instructional objectives can reflect objectivity. There is no system for collecting the comprehensive range of facts from which more valid inferences and objectives can be derived. Personal bias and power remain the chief determiners of educational objectives. This is true of the formulation of educational objectives at a national level, at an institutional level within a school system, and at the classroom level where the teacher selects objectives for individual pupils. Two indicators of possible change toward better sampling of pertinent information are (a) The Project TALENT Data Bank (1965) and (b) the National Assessment (McMorris, 1968). Project TALENT provides the first somewhat comprehensive and detailed study of the nation's youth that includes an inventory of their abilities in light of their personal and social needs and the social context in which they live. Similar data about crucial aspects of life now and in the future should also be made systematically available to curriculum planners.

## *Selecting Learning Opportunities*

Dressel and Mayhew (1954) concluded that objectives are best achieved when learning experiences are devoted to them. If an objective is to promote skill in problem solving, the learner must have an opportunity to solve problems. If the learner is to solve problems in mathematics, he must have an opportunity to solve problems in mathematics, not just practice solving problems in general or solving problems in other areas like health. The term "learning opportunities" means the situations, activities, objects or presentations which elicit desired responses from the learners. A book, film, or field trip that calls forth the particular responses desired is an example of a learning opportunity. The school administrator who hires a particular teacher for a given class of learners because that teacher will bring forth certain reactions from those pupils is engaged in selecting a learning experience, i.e., the teacher plus the responses desired.

Curriculum theory offers a number of general principles to guide one in selecting learning opportunities. One should select those situations that (a) give the learner opportunity to practice the objective or a prerequisite to the objective, (b) reward or reinforce the learner (e.g., experiences must be consistent with the learner's prior interest or predispositions), and (c) are in keeping with the learner's ability to respond. These general principles have been adopted from ancient pedagogical admonitions (Comenius, 1896) and psychological theories of learning (Hilgard, 1967). To the curriculum worker, however, these principles lack specificity. We cannot yet say how much opportunity to practice must be given a particular learner in attempting to help him accomplish a task. Neither can we state the degree to which the practice must be identical or equivalent to the objective. The curriculum worker also lacks knowledge of prerequisites for many of the instructional changes he seeks to bring about in the learner.

## *Instructional Variables and Research Strategies*

Programed instruction has been helpful in identifying prerequisite steps to attainment of objectives. Research in programed instruction has revealed reinforcement variables necessary for maintaining pupil responses and for indicating effective ways to elicit desired responses (Guggenheim, 1965). Theory and intuition are sources to use in selecting variables for manipulation. Markle (1965), for instance, developed a treatise dealing with generalizations for eliciting correct and relevant responses. Her rules for prompting and her cautions against the practice of having learners copy responses are logically derived; however, she puts final verification for instructional effectiveness upon empirical testing and has derived many of her principles after examining data from program tryout. Schramm (1964) listed nearly 190 original reports of research in programed instruc-



tion, most of which cast light upon requirements for selecting learning experiences. Among other things, he suggested the kinds of learning task that demand a constructed response; the beneficial consequences of gradually increasing presentation before practice; the importance of permitting students to select their own "size of step" (amount of presentation of information before practice); the need for making each response crucial, not merely trivial or easy; the conditions under which learners must or need not have knowledge of results to learn effectively; and how prompting can enhance results, particularly when there is a short time for practice.

Another excellent summary of experimental investigations of instructional variables was prepared by Lumsdaine and May (1965). In this review, the authors classified and related presentational variables, learner characteristics, and instructional objectives (task variables). Along the same lines, it has been suggested that researchers should attend more to "disordinal interactions" (Cronbach, 1957; Cronbach and Gleser, 1965). That is, they should seek to find instructional treatments that produce superior results with one type of learner but inferior results with learners of different characteristics. It is suspected (but not yet demonstrated) that learners who do not achieve by one treatment will achieve by a second treatment and conversely that there are those who would have achieved by the first treatment but not the second. Lumsdaine and May (1965) also mentioned a trend toward product testing in which one identifies what a particular instructional program will predictably accomplish when used in a given way with particular kinds of students.

The review by Lumsdaine and May treats over 100 experimental studies in some detail, thereby testifying to the popularity of current curriculum research upon the *how* (not the *what*) of instruction. There is noteworthy evidence that "programed instruction" has been broadened to include learning activities in a wide variety of media. Studies show that curriculum research in the area of selection of learning opportunities is characterized by progressive tryout and revision based on test data from preliminary versions of an instructional program. The studies reported differ from the curriculum practices of ten years ago when "gross methods" comparisons predominated. Instead of trying to show effects of "team teaching," "non-graded classrooms," "television" and "conventional instruction," researchers are attempting to manipulate narrowly, operationally defined variables, e.g., fixed versus student-regulated pacing (Briggs, 1961).

Accompanying the advance toward more operational definitions of instructional variables are different designs for conducting curriculum research in school settings. For example, new answers are being given for the problems of determining the appropriate experimental unit and unit of analysis (Campbell, 1963; Lumsdaine, 1963; Glass, Hopkins, and Peckham, 1969; Fletcher, 1968). Rath (1967) summarized the logical



and empirical defense for the argument that the proper unit of analysis in much school experimentation is the classroom and not the individual. For instance, in studies comparing one method with another in which all the children in one classroom receive one method, the appropriate unit of analysis must be the class and not the individual student. Thus, in conducting comparisons of five classes (30 pupils each) with another five classes (30 pupils each), one should conduct the statistical analysis by contrasting five versus five units—not 150 versus 150. Consequently, there should be more within-classroom experimentation in which treatments are presented to randomly selected individuals and not to the class as a whole. Experimentation within the classroom offers better control over time of day, the composition of the dynamics of the classroom and other elements which could influence results. A second potential force is Popham's *Simplified Designs for School Research* (1967), a programed booklet in which the author teaches how to distinguish between powerful and weak designs, and how to decide which of four designs are most suitable for different types of investigation.

Better definitions of variables and analyses of their interaction is undertaken to reduce the amount of trial and error in empirical tryout of curriculum materials. The search for dependable generalizations for constructing curriculum is a central task for the curriculum researcher. The final determination of valid principles is a long-range goal. Until that goal is reached, the best research strategy may be to rely on empirical demonstrations of the performance characteristics of particular programs.

### Teaching

In addition to the curriculum researcher's current emphasis on the pertinent variables in developing instructional materials, he is greatly concerned with describing teaching so he can understand and gain greater control over it. Macdonald (1965) found it useful to regard teaching and curriculum as two separate but congruent subsystems in which teaching is the organized presentation of stimuli and curriculum is the major source of stimuli found in instructional settings. In 1960, Bellack and Huebner wrote the first review of "teaching" as a phenomenon for curriculum investigation. They proposed that teaching behavior be studied in a context of theory and contended that the then prevailing modes of inquiry into teaching were not very fruitful. The common practice was to look for one or more of the following: (a) a relationship between teacher and pupil, (b) a description of a relationship between method and style of teaching, and (c) a conception of teaching (e.g., how teaching is like therapy). Bellack and Huebner criticized the lack of a theoretical framework which made it difficult to organize and to compare the sixty-three studies reviewed. Studies of teaching were, therefore, not cumulative, but isolated; and no study was built upon a predecessor. *The Handbook of*

*Research on Teaching* (Gage, 1963) presented a critical analysis of the literature on teaching which included sections on theoretical orientation. Callaway and Gleason (1963) reviewed studies of teaching within a rationale of curriculum. Micromatic approaches were singled out as attempts to develop a theoretical base for the teaching act. Flanders's method for analyzing teacher behavior (1961) and Smith and Meux's (1960) effort to examine the logical structure of the teacher's verbal behavior were given as examples of progress toward an understanding of teaching and the eventual development of a theory of instruction. These studies have been influential in directing curriculum activity towards methods for studying teaching. For instance, Bellack et al. (1963, 1965) extended the work of Smith by completing two related research projects entitled *The Language of the Classroom*. Bellack describes how to analyze classroom discourse to reveal the nature of the teaching "game." The rules of the game for pupils and teachers are derived from the analyses of actual classroom discourse using an original scheme for categorizing the verbal behavior of pupils and teachers. Bellack showed in these studies that it is possible to identify pedagogical moves and teaching cycles common to instruction. The variety of outcomes associated with classroom activity thus described is the subject of Bellack's current research. Gordon et al. (1963) moved more directly to the question of expected outcomes and teaching style. As described in *Dimensions of Leadership Behavior*, Gordon et al. (1963) conducted a study derived from both sociological and teaching theory; various modes of pupil morale, compliance, academic achievement and decorum were determined and operational generalizations were derived from the teaching style which maximizes desirable outcomes.

### *Concept of Mastery*

A final force to be reviewed, one of great potential for influencing views and practices in the selection of learning opportunities, is Bloom's (1968) concept of "Learning for Mastery." Borrowing the most promising ideas from many such as Carrol, Skinner, and Glaser, Bloom demonstrated a model which clearly indicates that if students are normally distributed with respect to aptitude and if the kind and quality of instruction and the amount of time available for learning are made appropriate to the characteristics and needs of *each* student, the *majority* of students can be expected to achieve mastery of the subject. Bloom's own class on test theory demonstrated increased cognitive outcome resulting from the use of the strategy for mastery. Before the strategy was adopted, approximately 20 per cent of the students reached a level of mastery; 80 per cent reached this same level the first year the mastery strategy was used; and by the third year, 90 per cent of a third class of students had achieved mastery. Among the key features of this mastery concept and strategy are the following:

- (a) Most students can master what is taught to them.
- (b) The task of the instructor is to find ways to enable students to master the subject.
- (c) Given enough time, nearly all students can attain mastery.
- (d) The learner must understand the nature of the task he is to learn and the procedure he is to follow in learning it.
- (e) It may be profitable to provide alternative learning opportunities.
- (f) The teacher should provide feedback on the learner's particular errors and difficulties.
- (g) Frequent feedback to the learners and specific instruction is effective in helping the learner to achieve.
- (h) The teacher must find ways to alter the time individuals need for learning.
- (i) Formulation of specific objectives of the learning task is an important precondition of mastery.
- (j) It is useful to break a course or subject into small units of learning and to test at the end of each unit.
- (k) Student effort is increased when small groups of two or three students meet regularly for as long as an hour to review their test results and to help one another overcome the difficulties identified by means of the test.

### *Organizing Learning Opportunities*

Courses are organized into a school program. Projects, units and topics are organized into courses. Activities are ordered within projects, units and topics. All purportedly have a common relationship among the objectives, the content and the behavior sought in learners. Typical questions of organization are: How often shall we provide for review? Can a key idea in one field be used in another field to expand the applicability of the concept?

Different views are guiding activity in organization of the curriculum. Gagné (1967) presented a psychological view; he drew heavily from learning theory to propose eight types of learning and to suggest that if these learnings are to occur, they must be learned in a hierarchal manner because the condition that distinguishes one form of learning from another is its prerequisite. Gagné saw problem solving, for example, as a type of learning that requires the learning of principles which in turn requires the learning of concepts, etc. Hewett (1964) also borrowed from psychology, particularly from reinforcement theory and research, to develop a hierarchy of learning tasks which is useful in teaching disturbed children.

The idea that subject matter itself should determine the sequencing



of instruction is implied in Tyler's (1967a) analyses of strengths and weaknesses of current research in science education. He criticized science programs because they lacked sequence. Tyler stated: "Since science is a structural discipline, rather than a miscellaneous collection of facts and generalizations, grade placement can be undertaken only in terms of sequence and not in terms of individual topics. None of the investigations conducted empirically tests sequences in one or more of the sciences." Tyler's comment implies that a structural discipline has an inherent order for teaching and learning—a controversial issue. Elsewhere, Tyler (1967c) was more open-minded and called for experimental work in curriculum organization in sciences; he listed some proposed alternatives for testing, e.g., moving from phenomena familiar to the child to those never seen by him; beginning with structural aspects of biological organism, then moving to functions; beginning with abstract analysis, then moving to application.

To some, knowledge is a hierarchy of ideas in which acquisition of the more complex ones depends on the previous mastery of simpler ones. Thus, when a curriculum designer has in mind a set of ideas he wants students to acquire, he should ask systematically: "What must the student already know so that he can acquire the new knowledge?" It follows that the best way to construct a textbook or an instructional sequence is to begin at the end and work backwards; evidence in support of this view has been presented by Gagné (1962). The identification and description of competencies for learning has been termed *task analysis*. Ways of conducting a task analysis or the identification of prerequisites are not well established, but empirical and logical strategies are currently being tried (Belcastro, 1961; Gagné and Paradise, 1961; Miller, 1962; Detambel and Stolurrow, 1956).

Not all instructional objectives have sequentially dependent "learning sets" through which the student must proceed to reach successively more complex kinds of competence. Mager and McCann (1963) demonstrated the superiority of learners who controlled the instructional sequence for certain training tasks. These researchers also reported that learners' motivation increased with the amount of control they were allowed to exercise over the learning sequence. Wulff and Kraeling (1961) trained subjects to note important features of all parts of a mechanism before they received any training in assembling it. These subjects learned less than those who used a procedure in which the same features were pointed out while the assembly was demonstrated.

Curriculum workers in vocational education promise to accomplish much in solving the problem of how best to do justice to each discipline and, at the same time, help the student use appropriate tools to deal with the world. Starling (1964), for example, studied the possibility for integrating biological principles and occupational instruction in agriculture.



There are signs that employing school children in a variety of industries will soon become popular in communities other than those with a history of high dropout rates. The development of programs in which students work to learn and earn affords curriculum planners the opportunity to renew old ideals of relating vocational education to general education.

There is a need for better definition of organizational variables and their interactions. Just as it is dangerous to generalize from a study comparing a live teacher with one on television so it is difficult to demonstrate from single instances that a plan of organization such as core courses (combination of two or more subjects in a double period) is more effective than a single subject-matter plan of organization. The core course or the single-subject course plan used in a given study may not exemplify the best in both approaches; neither are likely to be presented in a way that will allow replication. The results of organizational studies too often apply only to the particular programs compared, not to the whole class of organizational plans. Often it is the difference in amount of time permitted under the different organizational plans rather than the organizational design itself that accounts for the results; this only proves what is apparent—"If one gives more instruction, students are likely to learn more."

During the early 1960's, notions about team teaching, large and small group instruction, flexible scheduling, and continuous pupil progress were advanced as being relevant to curriculum organization (Goodlad and Anderson, 1959; Trump, 1959). Most of the literature on these topics was not based on systematic research although Gibboney et al. (1963) were able to identify two studies which purportedly evidenced the strength of team teaching and the desirable consequences of offering an honors track program to students. In his monograph on curriculum theory, Beauchamp (1961) wrote about the best way to organize the curriculum; he included a description of seven different bases for sequence that have been employed. He concluded, however, that the questions of organization still have to be studied and that there is more useful knowledge available for theory building in the area of organization than in any other dimension of the curriculum.

### *Evaluating Curricula*

Two influential specialists in curriculum evaluation are Ralph W. Tyler and Lee J. Cronbach. Tyler directed one of the few classic studies in curriculum; his *The Eight Year Study* was one of the most comprehensive studies of education ever undertaken. Tyler's protégé, Cronbach, continues Tyler's leadership in evaluation research by indicating new directions and procedures. Among the newer ideas expressed by Cronbach (1964) are that (a) there is greater need to describe how instruction modifies the student's thinking and attitude than there is to assign a

total score to his performance, and (b) present techniques of investigation are wrong, and as a result there is almost no pertinent information about how well ideas acquired in school are used several years later. Cronbach argued for more measures of how long it takes to recapture originally learned knowledge.

Scriven (1967) was influenced by Cronbach's work, but he extended the conception of evaluation in his "The Methodology of Evaluation." He stressed the importance of evaluation in the assessment of the end products of the curriculum and in guiding those developing that product. Scriven provided incisive arguments against the objections that value judgments are essentially subjective and not scientific. He advocated finding the match between goals and course content, between goals and examination content, and between course and examination content. Among Scriven's contributions is his specification of criteria which should be used in evaluating a teaching instrument. Indirect effects on other students, effects on administrator, range of utility, and secondary costs of curriculum adoption are among the new considerations he introduced.

Work is continuing in the preparation of guidelines for assessing the quality of specific instructional programs and for evaluating programmed instruction as a general method. Those interested in the major issues and technical problems of this effort should read "Assessing the Effectiveness of Instructional Programs" in *Teaching Machines and Programmed Learning II* (Lumsdaine, 1965). In this chapter Lumsdaine reviewed approximately 100 research articles pertinent to the methodological problems of improving assessment practices.

Ferris (1962) was one of the first to offer solutions to the problems of evaluating new curricula which had to be assessed in light of both traditionally oriented measures of achievement and measures of the gains for which the new curricula were designed. As director of Curriculum Studies at the Educational Testing Service, Ferris was responsible for finding out how students who were enrolled in four curriculum projects in science performed on College Entrance Examinations and in other test situations. His problem was to prepare items that did more than measure a student's understanding of *what* scientists had learned and a collection of technological terms. He sought to measure how well the students comprehended *how* the scientist learned. Questions were designed to require the student to apply what he had learned in the course to new situations. Ferris also measured the changes in pupils' attitudes toward science.

Heath (1964) demonstrated how curriculum advances when testing specialists apply themselves to research in curriculum. In his account of an experimental study in which teacher and curriculum effects were measured by a cognitive preference test, Heath pointed the way to better curriculum experiments.

Lord's "Estimating Norms by Item Sampling" (1962) is having an

impact on research and practice in evaluating the curriculum. Lord's study provided evidence that in some cases a better estimate of achievement (norm distribution) can be obtained by testing, say, 1,000 individuals in the norm population with just seven items (randomly drawn from seventy items) than can be obtained by giving seventy items to each person in a random sample of 100 individuals. As a consequence, curriculum evaluators are learning more about the various possibilities of trading items for people. Husek (1968) has given researchers a conceptual tool to make the assessment of instruction more efficient. His contribution is a series of views about testing purposes and kinds of testing instruments needed. He showed, for instance, the difference required in tests when the purposes are to evaluate students and to find out whether the instruction has been effective. Glaser (1963) has presented the concept of criterion-referenced measures. In accordance with this concept, the teacher assesses student achievement by using a criterion standard that indicates a student's degree of competence, independent of reference to the performance of others.

Not all researchers agree upon the methods and goals of curriculum research in evaluation. Atkin (1963) and Walbesser (1963) are representative of two different positions. Atkin believes that the testing for learning of certain pervasive concepts may itself be destructive. It may be, for example, that testing makes the student too self-conscious about the role of a concept at too early a stage. Walbesser's position is that it is less important to examine what a teacher or a curriculum writer says he himself is doing; the most important datum is what the student says (does, believes, etc.) at the end of the course that he wouldn't have said if he had not taken the course. The problems of evaluation in curriculum are receiving considerable attention. There is a healthy rapprochement between the test makers and curriculum developers which is long overdue.

### *Concluding Comment*

Today, curriculum as a field of study is attracting a wider range of participants. Philosophers, psychologists, social scientists, instructional technologists, measurement specialists, and pedagogues are laboring in the curriculum vineyard. It is doubtful if many working in curriculum really see their efforts as part of a whole, i.e., forces advancing the subject matter of curriculum. These participants are using newer tools to influence the conduct of curriculum such as criterion-referenced tests, task analysis, taxonomies of objectives and analytical schemes for describing teaching. Most of these tools are associated with a concern for accountability. They are instruments for checking out the results shown in learner behavior of teaching procedures, educational innovations and instructional plans and materials. Curriculum is becoming more rational because researchers and practitioners are beginning to realize that the desired changes in the learner are the true "ends" and that methods and instructional sequences used to produce those changes are the "means," not to be prized but appraised.



A balance exists between those whose sights are set on adapting curricula to critical social problems of the moment and those who are preoccupied with inventing curricula to serve the future. There is now increased concern with deriving and formulating objectives from observation of what is happening in the larger society. This concern has not, however, noticeably reduced interest in formulating objectives on the basis of organized knowledge (disciplines). There is wide interpretation and confusion of what is meant by formulating objectives in the light of data from learners.

More work is being undertaken on selecting learning opportunities (means) than on determining the long-range ends. Experimental procedures are gaining approval as the appropriate way to determine the value of learning opportunities. Although experimental procedures are applicable to problems of curriculum organization, few experimental research designs have employed them. Unfortunately, the general study of curriculum organization is neglected. However, curriculum evaluation is alive with new ways of thinking and working; there are many indications that curriculum and measurement specialists are beginning to work together.

It is encouraging to note the new emphasis upon experimental work *within* the field of curriculum. By manipulating variables in school practice and noting the consequences, curriculum knowledge is gained, and at the same time the objectives, content, and methods of instruction are changed. As long as studies in curriculum were descriptive and analytical, it was difficult for curricular knowledge to be a force in shaping what was taught and how it was taught. As I have indicated early in this chapter, there is an unfortunate lack of experimental studies showing cause and effect relations between social forces and curriculum change. Investigators have not manipulated what they consider to be forces influencing the curriculum and measured predicted consequences. Although it may be difficult to manipulate a societal force (e.g., a pressure group) someone should try it. If researchers and practitioners cannot exercise control over a force and regulate it in the interest of better curriculum, then why bother to identify it as a force?

### Bibliography

- American Educational Research Association. Curriculum Planning and Development. *Review of Educational Research* 36: 339-398; 1966.
- American Educational Research Association. Curriculum Planning and Development. *Review of Educational Research* 33: 227-337; 1963.
- American Educational Research Association. Curriculum Planning and Development. *Review of Educational Research* 30: 181-279; 1960.
- Ammons, Margaret. An Empirical Study of Process and Product in Curriculum Development. *Journal of Educational Research* 57:451-457; 1964.
- Anderson, C. C. Galbraith, Technology, and Education. *Alberta Journal of Educational Research* 14: 5-14; 1968.
- ASCD Commission on Current Curriculum Developments. *Influences in Curriculum*



- Change. Washington, D. C.: Association for Supervision and Curriculum Development, 1968.
- Association for Supervision and Curriculum Development. *Curriculum Change: Direction and Process*. Washington, D. C.: the Association, 1966.
- Atkin, J. Myron. Some Evaluation Problems in a Course Content Improvement Project. *Journal of Research in Science Teaching* 1: 129-132; 1963.
- Beauchamp, George A. *Curriculum Theory*. Wilmette, Ill.: Kagg Press, 1961.
- Belcastro, F. P. *Programmed Learning: Relative Effectiveness of Four Techniques of Programming the Addition and Subtraction Axioms of Algebra*. Doctor's thesis. Pittsburgh, Pa.: Univ. of Pittsburgh, 1961.
- Bellack, Arno et al. *The Language of the Classroom Part 1*. Cooperative Research Project No. 1497. New York: Columbia Univ., Teachers College, 1963.
- Bellack, Arno et al. *The Language of the Classroom Part 2*. Cooperative Research Project No. 2023. New York: Columbia Univ., Teachers College, 1965.
- Bellack, Arno and Huebner, Dwayne. Teaching. *Review of Educational Research* 30: 246-257; 1960.
- Black, Hillel. *The American Schoolbook*. New York: Morrow, 1967.
- Bloom, Benjamin S. Learning for Mastery. *UCLA-CSEIP Evaluation Comment* 1, May 1968. (Center for the Study of Evaluation of Instructional Programs, 145 Moore Hall, 405 Hilgard, Los Angeles, California).
- Bloom, Benjamin S. *Stability and Change in Human Characteristics*. New York: John Wiley & Sons, 1964.
- Bloom, Benjamin S. (editor). *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain*. New York: Longmans Green, 1956.
- Bobbitt, Franklin. *The Curriculum*. Boston: Houghton Mifflin, 1918.
- Borton, Terry. What Turns Kids On? *Saturday Review* 50: 72-74, 80; April 15, 1967.
- Brickell, Henry M. *Organizing New York State for Educational Change*. Albany: Univ. of the State of N.Y., 1961.
- Bricker, William A. Shaping the Behavior of Students. *Influences in Curriculum Change*. Washington, D. C.: Association for Supervision and Curriculum Development, 1968. Pp. 13-20.
- Briggs, L. J. Self-Pacing versus Automatic Pacing in Paired-Associate Learning. *Student Response in Programmed Instruction* (Edited by A. A. Lumsdaine.) Washington, D. C.: National Research Council. National Academy of Science, 1961.
- Broudy, Harry S.; Smith, B. Othanel; and Burnett, Joe R. *Democracy and Excellence in American Secondary Education: A Study in Curriculum Theory*. Chicago: Rand McNally, 1964.
- Bruner, Jerome S. Culture, Politics, and Pedagogy. *Saturday Review* 51: 69-72, 89, 90; May 18, 1968.
- Bruner, Jerome S. *The Process of Education*. Cambridge, Mass.: Harvard Univ. Press, 1960.
- Callaway, Rolland and Gleason, Gerald T. Teaching. *Review of Educational Research* 33: 293-300; 1963.
- Campbell, D. T. and Stanley, J. C. Experimental and Quasi-experimental Designs in Research on Teaching. *Handbook of Research on Teaching* (Edited by N. L. Gage.) Chicago: Rand-McNally, 1963. Chapter 5.
- Cawelti, Gordon. Innovative Practices in High Schools: Who Does What—and Why—and How. *Nation's Schools* 79: 56-88; April 1967.
- Center for the Advanced Study of Educational Administration. *Change Processes in the Public Schools*. Eugene, Ore.: Univ. of Ore., 1965.
- Coleman, James S. et al. *Equality of Educational Opportunity*. Washington, D. C.: U.S. Government Printing Office, Supt. of Documents, Catalog #FS 5.238: 38000, 1966.
- Comenius, John Amos. *The Great Didactic*. (Translated by M. W. Kettinge. London: Adam and Charles Black, 1896.
- Conant, James B. *The Comprehensive High School*. New York: McGraw-Hill, 1967.
- Conant, James B. *The American High School Today*. New York: McGraw-Hill, 1959.
- Cronbach, Lee J. The Psychological Background for Curriculum Experimentation. *Modern Viewpoints in the Curriculum*. New York: McGraw-Hill, 1964.
- Cronbach, Lee J. and Gleser, Goldine C. *Psychological Tests and Personnel Decisions*. Urbana: Univ. of Ill., 1965.

- Cronbach, Lee J. The Two Disciplines of Scientific Psychology. *The American Psychologist* 12: 671-684; 1957.
- Detambel, M. H. and Stolurow, L. M. Stimulus Sequence and Concept Learning. *Journal of Experimental Psychology* 51: 31-40; 1956.
- Doll, Ronald C. *Curriculum Improvement: Decision-Making and Process*. Boston: Allyn and Bacon, 1964.
- Dressell, P. L. and Mayhew, L. B. *General Education: Explorations in Evaluation*. Washington, D. C.: American Council on Education, 1954.
- Edling, Jack V. Educational Objectives and Educational Media. *Review of Educational Research* 38: 177-194; 1968.
- Edward A. Uhrig Foundation. *The Nature of Knowledge: Implications for the Education of Teachers*. Milwaukee: Univ. of Wis., 1961.
- Eisner, P. W. Educational Objectives: Help or Hindrance? *School Review* 75: 250-266; 1967.
- Fantini, Mario D. Taking Advantage of the Disadvantaged. *Teachers College Record* 69: 103-114; 1967.
- Ferris, Frederick L., Jr. Testing in the New Curriculums: Numerology, "Tyranny," or Common Sense? *School Review* 70: 112-131; 1962.
- Flanagan, John C. et al. *Studies of the American High School Project TALENT*. Monograph No. 2. Pittsburgh, Pa.: Univ. of Pittsburgh, 1962.
- Flanders, Ned A. Analyzing Teacher Behavior as Part of the Teaching-Learning Process. *Educational Leadership* 19: 173-180; 1961.
- Fletcher, Harold J. Possible Interpretive Problems in Analyses Using Group Means as the Experimental Unit. *Psychological Bulletin* 69: 157-160; 1968.
- Ford, G. W. and Pugno, Lawrence. *The Structure of Knowledge and the Curriculum*. Chicago: Rand McNally, 1964.
- Foshay, Arthur W. *Shaping Curriculum: The Decade Ahead. Influences in Curriculum Change*. Washington, D.C.: Association for Supervision and Curriculum Development, 1968. Pp. 3-12.
- Franklin, John Hope. Rediscovering Black America: A Historical Roundup. *New York Times Book Review*: 48-52; September 8, 1968.
- Gage, N. L. (editor). *Handbook of Research on Teaching*. Chicago: Rand McNally, 1963.
- Gagné, Robert M. The Analysis of Instructional Objectives for the Design of Instruction. *Teaching Machines and Programmed Learning, II*. Washington, D.C.: National Education Association, 1965.
- Gagné, Robert M. *The Conditions of Learning*. New York: Holt, Rinehart and Winston, 1967.
- Gagné, Robert M. and Paradise, Noel E. *Abilities and Learning Sets in Knowledge Acquisition*. Psychological Monographs: General and Applied, Vol. 75, No. 14 (Whole No. 518). Washington, D.C.: American Psychological Association, 1961.
- Gagné, Robert M. et al. *Factors in Acquiring Knowledge of a Mathematical Task*. Psychological Monographs: General and Applied, Vol. 76, No. 7 (Whole No. 526). Washington, D.C.: American Psychological Association, 1962.
- Galbraith, J. K. *The New Industrial State*. Boston: Houghton Mifflin, 1967.
- Gibboney, Richard A. et al. Curriculum Components and Organization. *Review of Educational Research* 33: 278-292; 1963.
- Gideonse, Hendrix D. The Relative Impact of Instructional Variables: The Policy Implications of Research. *Teachers College Record* 69: 625-640; 1968.
- Glaser, Robert. Instructional Technology and the Measurement of Learning Outcomes: Some Questions. *American Psychologist* 18: 519-521; 1963.
- Glass, Gene V; Hopkins, Kenneth D.; and Peckham, Perc D. *The Experimental Unit and the Unit of Statistical Analysis*. Research Paper No. 28. Boulder: Lab. of Educ. Research, Univ. of Colo. March, 1969.
- Godfrey, Eleanor P. *Audiovisual Media in the Public Schools, 1961-64*. Washington, D.C.: Bureau of Social Science Research, 1965.
- Goodlad, John I. and Anderson, Robert H. *The Non-Graded Elementary School*. New York: Harcourt, Brace, 1963.
- Goodlad, John I. and Richter, Maurice N., Jr. *The Development of a Conceptual System for Dealing with Problems of Curriculum and Instruction*. Cooperative Research Project No. 454. Los Angeles: Univ. of Calif. and Institute for Development of Educational Activities, 1966a.
- Goodlad, John I.; Steophasius, Renata; and Klein, M. Frances. *The Changing School*

- Curriculum. New York: The Fund for the Advancement of Education (477 Madison Ave., New York, N.Y. 10022.) 1966b.
- Gordon, C. W.; Adler, L.; and McNeil, John D. *Dimensions of Teacher Leadership in Classroom Social Systems*. Cooperative Research Project No. 1084. Los Angeles: Univ. of Calif., 1963.
- Guggenheim, Fred. Curriculum Implications and Applications of Programmed Instruction. *School Review* 73: 55-65; 1965.
- Heath, Robert W. Curriculum, Cognition, and Educational Measurement. *Educational and Psychological Measurement* 24: 239-252; 1964.
- Hechinger, Fred M. A Storm Gathers Across the Land. *New York Times This Week in Review*: September 4, 1968, Section 4.
- Hewitt, Frank M. A Hierarchy of Educational Tasks for Children with Learning Disorders. *Exceptional Children* 31: 207-214; 1964.
- Hilgard, Ernest R. Some "Principles" Potentially Useful in Practice. *A Basic Reference Shelf on Learning Theory*. Palo Alto, Calif.: Stanford Univ. and ERIC Clearinghouse on Educational Media and Technology, 1967.
- Hirsch, Werner Z. et al. *Inventing Education for the Future*. San Francisco: Chandler, 1967.
- Husek, T. R. The Implications of Different Types of Evaluation for Test Development. *The High School Journal* 51: 268-274; 1968.
- Husen, Thorsten. Curriculum Research in Sweden. *Educational Research* 7: 165-185; 1965.
- King, Arthur R., Jr. and Brownell, John A. *The Curriculum and the Disciplines of Knowledge*. New York: John Wiley & Sons, 1966.
- Kliebard, Herbert M. Curricular Objectives and Evaluation: A Reassessment. *The High School Journal* 51: 241-247; 1968.
- Krathwohl, David R. et al. *Taxonomy of Educational Objectives, Handbook II: Affective Domain*. New York: David McKay, 1964.
- Lord, Frederic M. Estimating Norms by Item Sampling. *Educational and Psychological Measurement* 22: 259-267; 1962.
- Lumsdaine, A. A. Assessing the Effectiveness of Instructional Programs. *Teaching Machines and Programmed Learning II*. (Edited by Robert Glaser.) Washington, D.C.: National Education Association, 1965.
- Lumsdaine, A. A. Instruments and Media of Instruction. *Handbook of Research on Teaching*. (Edited by N. L. Gage.) Chicago: Rand-McNally, 1963. Pp. 656-659.
- Lumsdaine, A. A. and May, Mark A. Mass Communication and Educational Media. *Annual Review of Psychology* 16: 475-534; 1965.
- Macdonald, James B. Educational Models for Instruction—Introduction. *Theories of Instruction*. (Edited by James B. Macdonald and Robert R. Leeper.) Washington, D.C.: Association for Supervision and Curriculum Development, 1965. Pp. 1-7.
- McGuire, Frederick and Kersh, R. C. *An Experimental Evaluation of Driver Education*. Berkeley: Univ. of Calif. Press, in press.
- McMorris, Robert F. National Assessment: Coming in 1968-69. *Phi Delta Kappan* 49: 599-600; 1968.
- Mager, Robert F. *Preparing Objectives for Programmed Instruction*. San Francisco: Fearon, 1961.
- Mager, Robert F. and McCann, John. *Learner-Controlled Instruction*. Palo Alto, Calif.: Varian Associates, 1963.
- Markle, Susan. *Good Frames and Bad*. New York: John Wiley and Sons, 1965.
- Matthews, Charles V. et al. *A Curriculum for Dropout-Prone Students—Delinquency Study and Youth Development Project*. Edwardsville, Ill.: So. Ill. Univ., 1966.
- Miller, Robert B. Task Description and Analysis. *Psychological Principles in System Development*. New York: Holt, Rinehart, and Winston, 1962.
- Oettinger, Anthony G. The Myths of Educational Technology. *Saturday Review* 51: 76-77, 91; May 18, 1968.
- Paulson, Casper F. Specifying Behavioral Objectives. *National Research Training Institute Manual for Participants in Research Development (CORD) Projects*. Monmouth, Ore.: Teaching Research Division, Ore. State System of Higher Education, August 1967. Part II, pp. 1-13.
- Phenix, Philip H. *Realms of Meaning*. New York: McGraw-Hill, 1964.
- Phi Delta Kappa. *Education and the Structure of Knowledge*. Chicago: Rand McNally, 1964.



- Phipps, Lloyd J. and Evans, Rupert N. Curriculum Development. *Review of Educational Research* 38: 367-381; 1968.
- Phipps, Lloyd J. and Fuller, Gerald R. *Project REDY, Rural Education-Disadvantaged Youth, A Research Prospectus*. Urbana: Agricultural Education Division, Univ. of Ill., 1965.
- Plowden, Bridget. *Children and Their Primary Schools: A Report of the Central Advisory Council for Education*. London: Her Majesty's Stationery Office, 1967.
- Popham, W. James. *Simplified Designs for School Research*. Inglewood, Calif.: Southwest Regional Laboratory for Educational Research and Development, 1967.
- Popham, W. James and Baker, Eva. Curriculum Principles for Prospective Teachers. *Teacher Education Quarterly* 22: 38-41; Spring 1965.
- Project TALENT Data Bank. Pittsburgh: Project TALENT, March 1965.
- Raths, James. The Appropriate Experimental Unit. Research in Review. *Educational Leadership* 25: 263-266; 1967.
- Research and Policy Committee of the Committee for Economic Development. *Innovation in Education: New Directions for the American School*. New York: the Committee, 1968.
- Rugg, Harold. The School Curriculum and the Drama of American Life. Curriculum-Making: Past and Present. Twenty-sixth Yearbook, National Society for the Study of Education. Chicago: Univ. of Chicago Press, 1926.
- Schramm, Wilbur. *The Research on Programmed Instruction*. U.S. Dept. of Health, Education and Welfare, Office of Education OE-34034, No. 35. Washington, D. C.: Supt. of Documents, Government Printing Office, 1964.
- Scriven, Michael. The Methodology of Evaluation. *Perspectives of Curriculum Evaluation*. (Edited by Robert E. Stake.) AERA Monograph Series on Curriculum Evaluation No. 1. Chicago: Rand McNally, 1967.
- Shaw, Frederick. The Changing Curriculum. *Review of Educational Research* 36: 343-352; 1966.
- Simpson, Elizabeth J. The Classification of Educational Objectives, Psychomotor Domain. *Illinois Teacher of Home Economics* 10: 110-144; 1966-67.
- Sizer, Theodore R. *Secondary Schools at the Turn of the Century*. New Haven, Conn.: Yale Univ. Press, 1964.
- Smith, B. Othanel and Meux, Milton O. *A Study of the Logic of Teaching*. Urbana, Ill.: Univ. of Ill., 1960.
- Smith, B. Othanel; Stanley, William O.; and Shores, J. Harlan. *Fundamentals of Curriculum Development*. New York: World Book, 1950.
- Smith, Melvin I. (director). *Teaching to Specific Objectives*. Second Progress Report for the Development of County Wide Articulation in Foreign Language Instruction Through Common Measurement Procedures. Modesto, Calif.: Stanislaus County Schools Office, 1968.
- Starling, John T. *Integrating Biological Principles and Instruction in Vocational Agriculture*. Columbus: Dept. of Agricultural Education, Ohio State Univ., 1964.
- Talmage, Harriet. *An Experimental Approach to Curriculum Engineering*. Doctor's thesis. Chicago, Ill.: Northwestern Univ., 1967.
- Torkelson, Gerald M. and Driscoll, John P. Utilization and Management of Learning Resources. *Review of Educational Research* 38: 129-159; 1968.
- Trump, J. Lloyd. *Images of the Future: A New Approach to the Secondary School*. Urbana, Ill.: Commission on the Experimental Study of Utilization of the Staff in the Secondary School, 1959.
- Tyler, Ralph W. Analysis of Strengths and Weaknesses in Current Research in Science Education. *Journal of Research in Science Teaching* 5: 52-63; 1967a.
- Tyler, Ralph W. A Talk with Ralph Tyler. *Phi Delta Kappan* 49: 75-77; 1967b.
- Tyler, Ralph W. *Basic Principles of Curriculum and Instruction*. Chicago: Univ. of Chicago Press, 1950.
- Tyler, Ralph W. New Dimensions in Curriculum Development. *Phi Delta Kappan* 48: 25-28; 1966.
- Tyler, Ralph W. Resources, Models, and Theory in the Improvement of Research in Science Education. *Journal of Research in Science Teaching* 5: 43-51; 1967c.
- University of Chicago. *Toward Improved Curriculum Theory*. Chicago: Univ. of Chicago Press, 1950.



- Walbesser, H. H. Curriculum Evaluation by Means of Behavioral Objectives. *Journal of Research in Science Teaching* 1: 296-301; 1963.
- Wulff, J. J. and Kraeling, D. G. Familiarization Procedures Used as Adjuncts to Assembly-Task Training with a Demonstration Film. *Student Response in Programmed Instruction*. Washington, D. C.: National Research Council, 1961.

## AUTHOR

JOHN D. McNEIL Address: University of California at Los Angeles, California Title: Professor of Education Age: 38 Degrees: A.B., San Diego State College; M.A., San Diego State College; Ed.D., Columbia Univ. Specialization: Experimental Studies in Curriculum and Instruction.

### 3: CURRICULUM MATERIALS

W. JAMES POPHAM\*

*University of California, Los Angeles  
and Southwest Regional Laboratory  
for Educational Research and  
Development*

Attention given to research on curriculum material might be justified as a logical outgrowth of the two chapters on learning materials in the June 1966 issue of the *Review* devoted to the topic of curriculum. However, there is a reason more compelling than historical precedent to scrutinize research on curriculum materials. An examination of the curriculum reform movement in this country during the 1960's reveals that, without exception, those curriculum projects which had the most significant effects upon educational practice produced curriculum materials to implement their new curriculum scheme. To state it bluntly, it takes more than admonitions from curriculum seers, even if they are accompanied by polished curriculum guides, to alter the procedures of busy educational practitioners. The educational reformer who eloquently urges classroom teachers to change their practices may receive the accolades of the educational community, but the educational reformer who provides a set of usable curriculum materials for the teacher is more likely to modify what goes on in the classroom. This statement is not meant to suggest that curriculum materials are the only vehicles for bringing about educational change, since there are other factors involved. However, the enormous impact of curricular materials as change agents should not be underestimated, and a consideration of research relevant to curriculum materials is clearly in order.

The broadest interpretation of the phrase "curriculum materials" includes such educational devices as textbooks, audio visual equipment, and bulletin board displays. However, because rigorous research regarding curricular materials will normally be focused on those materials as some form of treatment variable, it seems appropriate to consider only those curriculum materials which are essentially *replicable*, i.e., materials which permit repeated investigations of their attributes and effects. Unless one

\*Dr. Susan M. Markle, University of Illinois, served as a consultant on the preparation of this chapter. Dr. Popham is also indebted to Miss Jennifer Grant, University of California, Los Angeles, who coordinated the initial survey of the research on which this chapter is based.

attempts to study replicable curriculum materials, as opposed to those which are essentially unreproducible, it is impossible to attain generalizable results. This conception of curriculum materials as replicable phenomena has great similarity to the conception of an *instructional product*, an *instructional program*, which Markle (1967, p. 104) defined as "a reproducible sequence of instructional events designed to produce a measurable and consistent effect on the behavior of each and every acceptable student." Note that this definition of an instructional program specifies reproducibility both of the instructional *stimuli* and of the *effects* of the student's interaction with those stimuli.

Curriculum materials which are designed to be reproducible permit the developer systematically to improve the materials over time by testing their effects upon the intended learners. The possibility of a widespread application of such a trial-revision strategy to curriculum materials reflects a relatively recent approach to the development of such materials. For a variety of reasons, research devoted to the form of curriculum materials based on this strategy promises to yield considerable educational dividends. What, then, is the difference between programmed instruction materials and curriculum materials which possess replicability of presentation and effects? Essentially, there is almost no difference. Programed instruction specialists have been increasingly willing to expand their conception of what is considered a "program"; thus, the boundaries of their province now include a number of procedures which, in the early days of programed instruction, would have been rejected as "non-programs."

Because of the focus on this form of curriculum materials, several relatively recent reviews of research can be profitably consulted. Anderson (1967) treated a number of investigations in his consideration of the technology of instruction in educational psychology. Several classes of research studies described later in this chapter were examined by Anderson in his evaluative review.

An insight into the current state of curriculum material research in Europe can be gained from the comprehensive report of a major educational conference in England. Although the focus of the conference was on programed instruction, a number of aspects of education and technology are treated in this collection (Undwin and Leedham, 1967).

The April 1968 issue of the *Review of Educational Research* is also an extremely useful document; the chapter by Briggs on learner variables and educational media is particularly helpful.

### *The Classification Problem*

One of the most difficult tasks for a reviewer of research on curriculum materials is deciding upon a reasonable classification scheme. In the period reviewed (roughly from the beginning of 1966 to the close of 1968), one can find an increasing number of investigations concerned with instruc-

tional materials; many of these investigations deal with programed instruction. The research reports range from extremely sophisticated, analytic experiments to the most prosaic types of "Method A versus Method B" investigation. There were many studies on *treatment variables*, that is, modifications in various forms of the curriculum materials themselves or the manner in which they were used. A few investigations were found which dealt with the *utilization* of those materials and with their subsequent evaluation. Some research was identified which treated the *revision* and *evaluation* procedures to be employed in improving curriculum materials. It is encouraging that there was some investigative activity in all of these problem areas. Each topic will be treated in some detail in subsequent paragraphs.

### *Demonstrations of Effectiveness*

Many investigations were reported which demonstrated the effectiveness of diverse forms of curriculum materials in bringing about learner behavior changes. For example, Schurdak (1967) reported the use of a computer to teach the essentials of a FORTRAN course.\* Allen and others (1966, 1967) described the use of the *WFF'N Proof* game to produce a 17.3 point IQ improvement for those using the game in contrast with a 9.2 improvement for an untreated group. Neidt (1967) demonstrated the effectiveness of instructional video tapes in the promotion of study skills in a university setting. Popham (1966) reported the use of video tape programs to modify the behavior of prospective secondary school teachers.

Hartley (1966) reported that in comparing programed instruction and conventional methods (in terms of achievement and efficiency), 41 per cent of the early programs were significantly more effective than traditional methods, 49 per cent equally effective, and 10 per cent significantly less effective. Nonetheless many programed versus conventional instruction studies continue to be reported. One supposes that this kind of research was beneficial in the early days of programed instructional materials, but there seems to be no justification for such inquiries now. Typical of such research is Greatsinger's (1968) recent investigation in which he tested the difference in achievement between two groups of sixth-grade students, one group taught by programed materials and the other taught by conventional methods. There was no significant difference detected with respect to achievement, but the programed group spent 51 per cent less time than the conventionally taught group. It is strongly recommended that in the future such investigations not be reported in the literature unless they are designed to test the effects of some hopefully impactful treatment variation. As vehicles for advancing the state of knowledge regarding instructional materials, comparative studies of "programed versus conventional instruction" are thoroughly impotent. It should be noted

\*For a critical review of the Schurdak report see Kopstein and Seidel (1967).



that many of the so-called programs used in these comparative investigations were nothing but primitive first drafts and did not even satisfy the replicability requirements of the Markle definition cited previously in this chapter.

For those investigators who still yearn to evaluate the effectiveness of given instructional *approaches* (e.g., computer assisted instruction, programmed text, conventional text) by comparing one version of each, Kopstein and Seidel's observation (1967, p. 414) is relevant: "Logically this is equivalent to evaluating the potato-holding capacity of a sack, a box, and a barrel. So long as we are not given their dimensions or volumes, any comparisons are patently meaningless." Of course, evaluating CAI versus conventional instruction is fully proper if one is gathering data for local decision-making. Such evaluation does not contribute to the growth of a discipline, however.

Equally unfortunate are the continued reports of cross-media experiments in which comparisons of presentation media are made, implying that these results can somehow be generalized beyond the specific subject matter treated. Illustrative of such research is the investigation reported by Calder (1967) in which five methods to teach the operation of the 16mm projector were compared. The five treatments were slides, 8mm single concept film plus slides, slides followed by 8mm single concept film, 8mm single concept film used with slide sequence, and written instructions and illustrations. The results of the investigation indicated no significant differences among the five methods on a performance test; but on a transfer of training test, the method using a film with a slide sequence was significantly worse at the .05 level. Does this mean that in the future development of curriculum materials one should not coordinate films with slide sequences? Obviously not. The idiosyncratic nature of the instructional task makes results of this kind of investigation relatively ungeneralizable. An added complication arises when one realizes that the same instructional task will undoubtedly be approached quite differently by different materials developers. Since there is still so little technology and so much artistry in the development of curriculum materials, such cross-media studies yield relatively little useful knowledge. Research would be more helpful if it were not so wedded to the particular instructional topic, but permitted some generalizations across presentation media. Such investigations were not found during the three-year period under review.

### *Treatment Variables*

The majority of the curriculum materials investigations reported in recent years have dwelt on specific treatment variables associated with the materials. These investigations include both the particular attributes of the materials themselves and the manner in which they are used after they have been developed.

### "Organizers"

One relatively easily manipulated variable for the material developer is the kind of introductory material which will be given to the learners prior to or after their use of the materials. Drawing upon the works of Ausubel and others, Grotelueschen (1968) investigated the effects of differentially structured introductory materials on learner performance of a concept attainment task. His evidence confirmed the predicted hypothesis that introductory materials can markedly facilitate the performance of both high ability and low ability learners in complex learning tasks. Frase (1968), however, reported that questions placed *after* paragraphs were more helpful to learners than frequent pre-treatment questions. It was concluded by the latter investigator that frequent post-questioning either shaped or elicited appropriate reading skills, but pre-questions interfered with prose structure.

Rothkopf (1966) has considered such materials as "mathemagenic" or material facilitating learning by a focus effect. Bauman et al. (1969) reported three studies which indicate that placing an "organizer" after a lesson has a facilitative effect greater than that of an "advanced organizer."

Tiemann (1968) reported an extensive investigation of the development and revision of a televised course in economics. One aspect of his investigation contrasted the performance of students who had been provided with general objectives prior to instruction with the performance of those who had been provided with specific objectives. The results of a fifty-item multiple choice examination given at midterm revealed no difference between the two groups, but performance on 25 of the 50 items repeated as part of the final examination disclosed high scores to be associated with the specific objectives treatment ( $p < .05$ ). Tiemann concluded that the association between the objectives and the nature of the examinations was not perceived until the midterm examination, but thereafter this perception guided those students who received the specific objectives.

### *Relevant Practice*

Common sense suggests that the learner should be provided with practice relevant to the desired terminal behaviors, but there have not been an abundance of investigations over the years to support this contention. Several recent studies bear upon the question of whether relevant practice facilitates learning. Gropper (1968) contrasted the effects of giving "actual" or "recognition" practice to students learning to assemble a simple electrical motor. Instruction was presented by video tape. There was no limit on the amount of time given to students to assemble the motor on the criterion test. Error rate for the recognition practice group was four per cent, but for the actual practice group it was only two per cent. More important, however, was a marked difference between the two groups on the time it took to complete the assembly. The actual practice group took 25 minutes

while the recognition practice group required 32 minutes, i.e., 30 per cent longer.

Johnson (1968) investigated the combined effects of prompts, relevant practice, and feedback on the beginning teacher's ability to observe classroom performance recorded on video tape. A video tape was developed of 21 brief scenes of pupil-teacher interaction. Each scene was accompanied by a booklet in which a question directed the viewer's attention to a specific part of the scene and in which a different multiple choice question was to be answered. The booklet also provided the correct answer. The experimental students read the question, viewed the scene, selected the answer that best described the behavior in the scene, and were given the correct answer after responding. The control students did not receive such practice, but only viewed the video tape scenes. The performance of the experimental group was significantly superior to that of the control group. The investigator concluded that the learner must be told what to focus on, given an opportunity to practice that behavior, and given corrective feedback about the accuracy of his response.

The value of relevant practice was also reported by Melching and Nelson (1966) in an investigation in which they provided practice frames in a programmed military training course on counter-insurgency. One group received programs including the practice frames; a second group received similar programs without the practice frames. The presence of the practice frames augmented achievement on a written recall test. The use of practice frames also apparently resulted in fewer program errors and faster progress through the program.

While investigating the importance of "problem heterogeneity" on instructional materials, Traub (1966) discovered that sixth-grade pupils who were permitted to practice on heterogeneous problems (as opposed to groups permitted to practice on homogeneous or irrelevant problems) performed significantly better on a complex mathematical task involving the solution of heterogeneous problems. Similarly, Williams (1966) concluded that achievement is enhanced when the correlation between training and type of criterion item is high. She used a group of sixth-grade pupils in studying the relationship between the response mode during training and the form of item provided in the final examination.

These five investigations, therefore, support the contention that curricular materials should be designed to provide the learner with opportunities to practice behaviors relevant to the desired terminal behavior.

### *Knowledge of Results*

In his review of research on knowledge of results, Anderson (1967) concluded that although knowledge of results does facilitate learning, there is considerable doubt that it functions primarily as a reinforcer. The pro-

vision of knowledge of results in instructional materials, he contended, although positive, must be studied further.

Karraker (1967) tested the hypothesis that plausible wrong responses in multiple-choice items are subsequently recalled as being correct. Seventy-two college freshmen were stratified according to high and low ability, then randomly assigned to *no-knowledge of results*, *knowledge of results*, or *control* treatments. The subjects responded on a multiple choice test in educational psychology; a recall criterion test was administered later. A subject's score was the number of responses that were the same as the plausible wrong responses in the multiple-choice test. The *no-knowledge-of-results* group made significantly ( $p < .01$ ) more errors than the *knowledge of results* or *control* groups. When no knowledge of results was given, subjects did recall more plausible wrong responses as being correct, but this effect did not appear when knowledge of results was given. The investigator concluded that when knowledge of results was given, plausible wrong responses had no more effect than when subjects were never exposed to them.

There is some suggestion that a variety of procedures can be used for supplying the learner with knowledge regarding the appropriateness of his responses, and that there may be very little difference among these feedback procedures. Meyer (1966) compared the effectiveness of four methods of confirming responses in a set of self-study curriculum materials. The topic was a refresher course in engineering for air force personnel. The methods of response confirmation were (a) chemically impregnated inks which changed color when moistened, (b) punchboards, (c) opaque erasable overlays, and (d) a machine with a lighting and buzzing system. No significant differences in learning were found to result from use of these methods, nor were there any notable differences among group opinions about the methods.

Attempting to analyze the impact of feedback on learners of different types, Campeau (1968) compared the performance of children classified as "high anxious" and "low anxious" on the basis of the *Test Anxiety Scale for Children*. Under conditions of knowledge of results and no knowledge of results in a learning program, an interaction was detected wherein highly anxious girls did better under knowledge of results conditions and low anxiety girls excelled under the no knowledge of results conditions. No significant main or interaction effects were detected for boys.

These various investigations are consonant with the bulk of previous research which supports the efficacy of using procedures in curriculum materials whereby knowledge of results may be secured.

### *Promoting Learner Interest*

Several investigators have reported studies in which deliberate attempts were made to devise curriculum materials which would be interesting



to the learners. While testing the effects of different motivational strategies with five-year-old disadvantaged learners using an automated teaching system, Gotkin et al. (1967) discovered that subjects made substantially fewer irrelevant responses in those sections of the instructional materials which incorporated motivational strategies (such as appealing pictures, stories, etc.) than in those sections without such devices. They also found that the total number of irrelevant responses was lowest during the early part of the instructional program; the investigators assumed that this finding was partially explained by the novelty of using the machine. The investigators observed that the need for motivating contexts became apparent as the frequency of irrelevant behavior increased with continued experience on the machine.

An investigation of context-based frames (with relevant supporting material) versus primitive copying frames without contexts (Faust and Anderson, 1967) revealed that in two experiments those learners exposed to the context program recalled significantly more ( $p < .05$ ) overtly practiced Russian words on posttests than did the no context groups. The advantage of the context learning materials was especially pronounced for subjects who hurried through the program.

Nagel (1968) reported a similar investigation in which it was demonstrated that a controversial and anecdotal style of writing, when compared with a formal style of writing, resulted in significantly greater "treatment effect" from a set of materials dealing with celestial navigation. Eighty naval reservists were randomly assigned to two versions of an 81-frame program. One program was written in a conversational style and the other was written in an extremely formal style. The investigator concluded that the type of writing used has significant effects on the performance of individuals who have no previous experience with programmed instruction or with the subject matter being studied. Unfortunately, in the latter investigation several important principles were violated which should be adhered to in conducting such research. His use of extremely short treatment times (only an 81 frame-program) makes it difficult to generalize to materials which will be used for the typically longer presentation times. The investigator attempted to question subjects about the "attitudes toward programmed instruction" on the basis of this single exposure. This is analogous to placing a child before a television set for the first time, letting him view the Lawrence Welk Show, and then asking him about his "attitude toward TV."

All three of these investigations, however, tend to support the generally accepted principle that interesting curricular materials will be more effective than uninteresting materials.

### Prompts

The degree to which prompts are useful in the design of curricular

materials has received attention recently from Richard C. Anderson and his associates. In one of these studies (Anderson and Faust, 1967) the value of underlining response terms in a programmed lesson as a means of increasing learning was investigated. Sixty-one graduate students received a training sequence involving ten presentations of 16 English-Russian word pairs and several tests. Two versions of the program were used. In one version the Russian word to be written in the blank was always underlined. In the second version, no words were underlined. The underlined version of the program was less effective than the version in which the subject was required to give more attention to the response.

In another investigation (Anderson et al., 1968), 108 subjects completed either a standard or a heavily prompted version of a 1,052 frame section of a psychology program. As predicted, those who received the standard program scored higher on the posttest than those who received the heavily prompted version. The results were interpreted as showing that learner performance is decreased by arrangements of curriculum materials which permit the student to respond correctly without noticing the important cues. The investigators inferred that the heavily prompted program resulted in poorer achievement than the standard version because it permitted attenuated inspection behavior.

### *Multiple Channels*

Travers (1967) has questioned the generally accepted belief that multimedia presentations are intrinsically more effective than single-channel presentations. Research dealing with this question continues. Severin (1967) investigated the effectiveness of relevant pictures in multiple channel communications. The study was cast in the context of a cue summation theoretical view which predicts that learning is increased as the number of available cues or stimuli is increased. Six message conditions were set up as follows: (a) an audio message only, tape recorded; (b) a print message only, recorded on 2 X 2 slides; (c) the audio and print messages combined for simultaneous presentation; (d) the audio with related pictures on 2 X 2 color slides for simultaneous presentation; (e) audio as in the first treatment with unrelated pictures of the same category, e.g., the word "moose" with a picture of a bison, on 2 X 2 color slides presented simultaneously; (f) the audio with unrelated pictures of a different category, e.g., the word "moose" with a picture of a catfish, on 2 x 2 color slides presented simultaneously. Two hundred seventh-grade students were stratified on intelligence and randomly assigned to one of the six treatments. The results were consistent with the prediction from the cue summation theory. When a series of words was tested for recognition, audio with related pictures was significantly better than audio accompanied by print. The combination of audio and print was not significantly better than print alone, which would be predicted by the absence of cue summation. Print was significantly better than audio, and print with audio was

significantly better than audio alone. Because of the irrelevant cues introduced, the combinations of audio with unrelated pictures were the poorest. The findings were in agreement with Travers's conclusion that a redundant multichannel communication is not significantly better than a single channel communication.

Kanner's review (1968) of the research on the instructional effectiveness of color in television led him to conclude that research consistently demonstrates that color does not provide important cues to learning not found in black and white presentations. Kanner suggests that verbal cues can substitute for the use of color with no loss in learning efficiency. The pleasure or satisfaction obtained by looking at color television was not disputed.

### *Sequencing*

Probably because it is relatively easy to alter a sequence of instructional events, the question of instructional sequence has been studied by many investigators. Gagné's (1965) work in this field has stimulated a number of researchers to investigate the sequence issue. Payne et al. (1967) conducted one of the more systematic investigations on this topic. They examined the effect of scrambling upon student learning both in terms of immediate achievement and delayed retention. Two hundred thirty-eight college sophomores were randomly assigned to eight groups to learn three basic concepts of educational measurement. The judges determined which of the concepts was most dependent upon logical development of content and each of the concepts was programed in both scrambled and logical form. After the eight groups had completed the three programs, an immediate test was given and a retention test was administered two weeks later. There was no confirmation of the hypothesis that the scrambling of frames would interfere most with the topics possessing the greatest logical development. There was also no substantiation for the hypothesis that students with high ability might be able to overcome the effect of scrambling more readily than those with low ability.

Several other investigators have compared the effects of scrambled versus logically ordered programs. Neidermeyer (1968), using one of Gagné's programs, conducted a study involving logical order, scrambled, and reverse order sequences. Although there was some evidence that the logical sequence was more efficient, the investigator concluded that sequence in short programs may not be as crucial to cognitive outcomes as has been previously thought.

Wodtke et al. (1968) described a similar investigation involving two different topics, one with a supposed conceptual hierarchy and one consisting of relatively discrete facts. Scrambled and logical sequences for both topics were developed and used with college students. Failures to detect differences between the performance of those exposed to scrambled

and logically organized materials were attributed to the ability of college students to intellectually reorganize the scrambled materials.

On the other hand, Biran (1967) reported opposite results from an investigation comparing scrambled and sequential versions of a branching program with sixth-grade children. The sequential version resulted in higher test scores, required less study time, and elicited fewer critical comments.

Allen and McDonald (1966) examined the value of allowing the student to determine his own organization of the instructional materials he encounters. Three versions of curriculum material organization were used with 65 students in an educational psychology class. In one of the versions students had complete control of the organization, in a second version students exercised partial control, in the third the instructor was completely in control. The major conclusion of their investigation was that performance was significantly lower when the student completely controlled the organization of the instructional materials and the method used in learning than when these elements were completely or partially controlled by the instructor. The difference in the time required to complete the instruction, however, favored the learner-controlled organization. The investigators concluded that the advantages of student control of instruction need to be carefully specified and tested, and that learner control of curricular sequence should not be assumed to be the most desirable scheme for organizing content.

Wendt (1967) conducted an experiment on a branching instructional program to discover what percentage of students would choose the option to return to the main line of instruction when given a choice between returning or skipping ahead. According to the investigator, the study indicated the need for *optional* returns from bypasses since a substantial percentage of the subjects (18 out of 41, or 44 per cent) refused the earned bypass at least once and 12 per cent refused every earned bypass. Some students, the investigator speculated, would definitely prefer to return to the main line of instruction rather than skip ahead as one might think.

Moore's (1968) investigation of three principles used in the preparation of curriculum materials, *viz.*, gap, irrelevancies and mastery, involved 184 eighth-grade students and a subject matter dealing with test-taking strategies. A  $2 \times 2 \times 2$  factorial design was used in studying each of the three principles and in testing the various main effect and interaction hypotheses. The investigator concluded that the most predictable finding was that introducing gaps into the instructional materials resulted in decreased learning performance. The most interesting results, according to the investigator, were the interaction effects of the mastery and relevancy principle. When material irrelevant to the objectives was introduced and mastery of it was required, subjects performed significantly worse. There was no discernible



detrimental effect when irrelevant material was presented but not required to be mastered.

An investigation by Harley and Woods (1968) failed to confirm the predicted effectiveness of backward chaining (commencing the sequence with elements of the final frames) as a sequencing strategy when used to teach poetry to a group of elementary school girls. The flurry of investigation regarding sequencing variables has not yet offered clear guidelines for the development of curricular materials, but it seems likely that continued research on this topic will soon yield more definitive prescriptions.

### *Pacing*

A number of investigators (e.g., Allan and Richardson, 1967) have conducted research on the rate at which learners should be exposed to curricular materials. Such research has, in some cases, gone beyond the group-paced versus self-paced studies which were fairly common in the early days of programmed instruction. Investigators are now attempting to find criteria for deciding between group- and self-pacing; and, after making this basic decision, how these forms of pacing can best be accomplished.

A series of experiments regarding group- versus self-pacing was conducted in the Dundee (Scotland) College of Education and reported by Allan and Richardson (1967). The bulk of the Dundee data tends to support the efficacy of group-paced instruction. The investigators speculated about the reason for these findings. They observed that because group-paced presentation is supposed to run counter to the individualized nature of programmed instruction, no program is constructed for an individual child. The programmer builds it for "the optimum child." The four major concerns of program builders are amount of information per frame, amount of illustration per frame, response quality, and rate of presentation; the first three are determined by "the optimum child." If programs were individualized, each one would be altered according to the individual. However, the programmer aims at the widest coverage consistent with these criteria. The Dundee investigators contend that group-paced programs, presented at the mean rate for the group, are also aimed at the "optimum child."

Sawiris (1966) reported an investigation of group-paced versus self-paced instruction for 124 English elementary school children on a set of curriculum materials dealing with geometric ratios and proportions which was consistent with the Dundee experiments. No significant differences between the two presentation rates were detected. The investigator concluded that, contrary to common belief, there were apparently some instances in which group-pacing could produce results comparable to self-paced materials. However, other English investigators (Fellows and Apter, 1967) reported investigations favoring individually rather than group-paced curriculum materials for young children.

Gallegos (1968) studied the effects of fast, slow, and self-pacing modes on high ability and low ability students who used programmed curriculum materials. Although the high ability students significantly out-performed their low ability counterparts in both fast-paced and self-paced treatments, no such differences were found under the slow-paced treatment. The research also did not support the contention that students, regardless of ability, can successfully proceed through programmed materials simultaneously at an externally established rate. All students achieved significantly inferior scores when proceeding under externally established fast-pacing conditions. Low ability students did better under slow external-pacing than under self-pacing. The investigator urged that more flexible pacing modes be developed to accommodate ability differences in students.

Probing the degree to which the learner can control his own pace, Kress and Gropper (1966) studied the self-pacing behavior of eighth-grade students as they completed two linear science programs. The time required to complete the program was found to be only moderately related to intelligence. Each student's work rate appeared to be determined by a number of factors besides intelligence. These investigators suggested that it would be desirable to control student learning behavior to a greater degree than is now being practiced. They contended that only those students who demonstrate on screening programs that they can work at a pace which permits them to reach satisfactory achievement levels should be permitted to pace themselves.

## *Revision and Evaluation*

### *Empirically Based Revision*

Recently (at an informal colloquium\*), Arthur A. Lumsdaine observed that there was little research which demonstrated that revision based on empirical tests, as opposed to skilled editorial revision, produced better learner achievement. Lumsdaine's observation is a challenge to those involved in the development of curriculum products. It is almost universally assumed that empirically-based tryouts and revisions yield better materials, but a survey of the literature reveals only a few studies which corroborate this position. Gropper et al. (1961) showed that revised materials—a lesson on the effects of heat and an introduction to chemistry—augmented learner performance from 6 to 26 per cent. Other than this classic study, few studies clearly support the value of revisions. Two studies which have not been reported in the journals but which tend to support the utility of revising (empirically) curricular materials were those conducted by Robeck (1965) and by Tiemann et al. (1966).

Gropper (1967) reported an investigation in which revised materials

\*Product Research Training Program Colloquium, University of California, Los Angeles, May 19, 1967.

based on pupil performance were used with over 100 seventh-grade pupils. Revised versions surpassed the original versions by approximately 30 per cent on immediate test scores, by approximately 20 per cent on retention test scores, and by approximately 40 per cent on transfer test scores. It was also noted, however, that the revised versions took approximately twice as long to complete as the original materials. (One can speculate about the type of impact that the generally rapid-paced sequence has on learners in the typical teacher's classroom.)

A series of investigations by Vander Geer and his associates (1964 a, b, c, 1965) offers some informal cues regarding variables which may be profitably manipulated in revision studies. Vander Geer's studies were related primarily to modifications of filmstrips and similar media.

Scriven's (1967) distinction between *formative* and *summative* evaluation is being used with increasing frequency in the literature of curriculum research. Formative evaluation concerns evaluating the curriculum materials so they can be improved while they are in the developmental stages. Summative evaluation concerns the evaluation of the materials (usually by contrasting them with other curriculum materials) after they have been developed. Dick (1968) dealt with several questions related to formative evaluation as he studied the behavior of lesson revisers. The lesson revisers were students from four colleges and universities using a specified calculus program. They were provided with the following data which might be used for revision of the program: (1) item analysis of the end-of-lesson test, (2) a listing of the incorrect answers to the test items, (3) a guide or reference sheet which indicated where in the programmed book the information relevant to each test item was taught, (4) the error rate for each of the program frames, (5) a sample of the incorrect responses to the frames, (6) student comments about various parts of the assignment, and (7) reviewer and instructor comments. The consensus among the revisers was that the frame error rates and the reviewer comments constituted the most informative data. Few of the revisers studied the results of end-of-lesson tests or attempted to relate test item performance to the particular frames in the program. End-of-lesson tests *not* constructed by the revisers themselves were viewed as not completely representative of the objectives the revisers would have tested.

If Dick's survey of untrained reviser behavior is typical, one realizes how necessary proper training is for those modifying curricular materials. The revisers in this study were simply attending to the wrong criteria. For example, Anderson and Faust (1967) demonstrated that error rate is highly dependent on framing technique, yet may be unrelated to criterion performance.

### *Evaluation*

One of the more interesting developments concerning the evaluation



of curriculum materials has been the establishment of the Educational Products Information Exchange (EPIE) Institute. EPIE is an independent, nonprofit organization that was chartered in 1967 to serve educational decision-makers in schools and industry by supplying information about the availability and effectiveness of instructional materials and equipment. The publications of EPIE since its establishment have focused largely on the evaluation of educational equipment (particularly of an audiovisual nature), but the institute plans to undertake systematic evaluations of educational curriculum materials. Thus far the majority of EPIE publications regarding curriculum materials has unfortunately been descriptive rather than evaluative. Since October 1968, the institute's journal (formerly called the *EPIE Forum*) is known as the *Educational Product Report*.<sup>\*</sup> It is reasonable to expect that if EPIE and similar organizations survive, they will provide educators with something akin to a *Consumers Report* for curriculum materials and related products. Such reports, if systematic, sensitive, and objective, should clearly be of considerable utility in choosing among competing curriculum materials.

Programed instruction has generally had a beneficial impact on the quality of student learning, but has provided measurement and evaluation personnel with a number of problems about the manner in which the attainment of educational outcomes should be assessed. Of particular interest are those investigations concerned with the difference between norm-referenced and criterion-referenced measurement approaches (see Schutz, Chapter 6, this issue). These approaches are being employed by material developers. For example, Newmark and Sweigert (1967) tested three different approaches to the teaching of Spanish for sixth-grade students: (1) instruction by a qualified foreign language teacher using the Modern Language Association language course, (2) programed self-instruction, and (3) instruction by television. The investigators concluded that their study demonstrated the feasibility of applying criterion-referenced testing to a large-scale field test of three foreign language courses. The study indicated that traditional, norm-referenced methods, when used to evaluate a small sample of the specific linguistic objectives of a language course, may obscure deficiencies in the material or learning conditions. The test results show that with few exceptions all three language courses failed to achieve their basic linguistic objectives under the learning conditions in the study.

### Utilization

Several investigators have been concerned with the problem of the proper installation of curriculum materials after they have been satisfactorily developed. Tobias (1966, 1968) studied the degree to which teachers are

<sup>\*</sup>The *Educational Product Report* is published nine times a year by the Educational Products Information Exchange Institute, 386 Park Avenue, South, New York, N. Y. 10016.



threatened by language regarding curriculum materials which connotes "automation" rather than comparable terms not identified with automation. The investigator observed, on the basis of samples of teacher's attitudes toward and knowledge about instructional technology, that many teachers think they are threatened both by newer media and by "machines" which might replace them. Their negative reaction appears to be only slightly affected by the amount of information about modern instructional technology that the teacher possesses. Tobias suggested that mechanistic terminology used in the description of new curriculum materials is not particularly useful and may arouse avoidable fears.

La Giapa (1968) investigated the impact of the ability level of teachers on the teaching of programmed instruction materials. On the basis of supervisor's ratings, the eight most competent and eight least competent teachers were identified and permitted to use conventional and programmed instructional approaches in classes for the military. For difficult material, it was found that the programmed materials were more effective with the weaker teachers ( $p < .05$ ); no such differences were noted in the case of easier materials. The investigator concluded that the degree of superiority of programmed instructional materials over conventional instruction is influenced both by the teacher's ability and the subject matter difficulty. The effects of programmed instruction, he continued, are apparently greater with low-ability than with high-ability teachers and with complex subject matter rather than simple.

From a methodological point of view, the more provocative research on the use of curriculum materials has been conducted by investigators who have used physiological measures, such as galvanic skin response apparatus, during the presentation of curricular materials. Levonian (1967, 1968) concentrated on assessing the degree of learner arousal during instructional sequences. In one of his investigations, for example, the results suggested that forgetting occurs when information is learned under conditions of low learner arousal, whereas reminiscence occurs if information is learned under conditions of high arousal.

#### *Availability of Presentation Devices*

Of some interest to those developing curriculum materials are investigations such as the survey conducted by the National Education Association Research Division (1967) in which a total of 1,609 elementary and secondary teachers across the nation answered questions related to instructional media. In preparing non-printed curriculum materials, the availability of instructional devices such as audio tape recorders is of significant interest to the developer. If it is true that less than 5 per cent of the nation's educators have access to computer-based teaching terminals, it would seem unwise for a curriculum developer to devote his efforts exclusively to the preparation of materials which require such terminals. Obviously the equipment situation may change from year to year, but one should recognize

that equipment availability does, to some extent, limit the form of potential curriculum materials. The results of the 1967 NEA survey indicated that all teachers (elementary and secondary combined) had access to the following equipment in the percentages indicated: phonograph, 93 per cent; silent film strip projector, 92 per cent; 16mm motion picture projector, 85 per cent; overhead projector, 83 per cent; audio tape recorder, 78 per cent; opaque projector, 73 per cent; sound film strip projector, 54 per cent; 8mm motion picture projector, 27 per cent; closed circuit television facility, 11 per cent; computer-based teaching terminal, 3 per cent.

### *Research Gaps*

From an inspection of the research related to curriculum materials during the past several years, one is impressed by several deficiencies. First, studies of the revision process to improve the quality of curriculum materials have not been frequently conducted. There seems to be an overriding faith in the idea that materials revised to be consistent with empirical tryout data will become better, but this has not been clearly demonstrated. Certainly the manner in which revisions can be made most efficiently has not been carefully treated.

Although there are some investigators who are beginning to study ways of evaluating the quality of curriculum materials with diverse kinds of measures (measures which deal with noncognitive as well as the more common kinds of cognitive learner outcomes), there is a great need for research on newer evaluative techniques. Can affective measures be devised to reliably assess the learner's approach tendencies toward particular kinds of curriculum materials? How can researchers determine the degree to which the learner will remain "plugged in" to the materials as he goes through them? Mager's (1968) text about measurement and instructional strategies related to learner interest is relevant. Evaluative research which attempts to study other than standard types of measurement strategies definitely needs to be conducted.

Some of the investigations reviewed in this chapter could easily be used by materials development personnel. The hope is that more studies will be conducted that possess clear relevance to the practical decisions which must be made regarding preparation of curriculum materials.

It was observed at the outset of this chapter that curriculum materials may provide one of the most formidable change agents for modifying educational practice. To the extent that this is true, educational researchers obviously need to devote more systematic attention to this area of research than they previously have.

### *Bibliography*

- Allan, J. I. and Richardson, R. H. *Programmed Learning—A Multi-Media Approach. Programmed Learning and Educational Technology* 4: 191-95; 1967.  
Allen, Dwight W. and McDonald, Frederick J. *The Effects of Self-Selection on Learning*

- in Programmed Instruction. *American Educational Research Journal* 3: 1-6; 1966.
- Allen, Layman E.; Allen, Robert W.; and Miller, James C. Programmed Games and the Learning of Problem-Solving Skills: The WFF'N Proof Example. *Journal of Educational Research* 60: 22-26; 1966.
- Anderson, Richard C. Educational Psychology. *Annual Review of Psychology* 18: 129-64; 1967.
- Anderson, Richard C. and Faust, Gerald W. The Effects of Strong Formal Prompts in Programmed Instruction. *American Educational Research Journal* 4: 345-52; 1967.
- Anderson, Richard C.; Faust, Gerald W.; and Roderick, M. C. Over-Prompting in Programmed Instruction. *Journal of Educational Psychology* 59: 88-93; 1968.
- Bauman, Daniel J.; Glass, Gene V.; and Harrington, Scott A. The Effects of the Position of an "Organizer" on Learning Meaningful Verbal Material. *Research Paper No. 24*. Boulder: Lab. of Educ. Res., Univ. of Colo., 1969. (Offset.)
- Biran, L. A. A Comparison of a "Scrambled" and a "Sequential" Presentation of a Branching Program. *Programmed Learning and Educational Technology* 4: 290-95; 1967.
- Briggs, Leslie J. Learning Variables and Educational Media. *Review of Educational Research* 38: 167-73; 1968.
- Calder, Clarence R., Jr. A Comparison of the Relative Effectiveness of Five Self-Instructional Techniques in Teaching a Psychomotor Skill. *Educational Technology* 7: 14-19; 1967.
- Campeau, Peggie L. Test Anxiety and Feedback in Programmed Instruction. *Journal of Educational Psychology* 59: 159-63; 1968.
- Dick, Walter. A Methodology for the Formative Evaluation of Instructional Materials. *Journal of Educational Measurement* 5: 99-102; 1968.
- Faust, Gerald W. and Anderson, Richard C. Effects of Incidental Material in a Programmed Russian Vocabulary Lesson. *Journal of Educational Psychology* 58: 3-10; 1967.
- Fellows, B. J. and Apter, M. J. An Experimental Study of Three Variables in the Presentation of Audiovisual Teaching Programs. *Programmed Learning and Educational Technology* 4: 19-28; 1967.
- Fraser, Lawrence T. Effect of Question Location, Pacing and Mode Upon Retention of Prose Material. *Journal of Educational Psychology* 59: 244-49; 1968.
- Gagné, Robert M. *The Conditions of Learning*. New York: Holt, Rinehart and Winston, 1965.
- Gallegos, Arnold M. Some Evidence and Ideas About Pacing and Ability. *National Society for Programmed Instruction Journal* 7: 12-14; 1968.
- Gotkin, Lassar G.; McSweeney, Joe; and Richardson, Ellis. Acquisition of Pre-Reading Skills in Five-Year-Old Children from Disadvantaged Backgrounds Using the E.R.E. Automated Teaching System. *National Society for Programmed Instruction Journal* 6: 7-9, 18; 1967.
- Greetsinger, Calvin. An Experimental Study of Programmed Instruction in Division of Fractions. *AV Communication Review* 16: 57-62; 1968.
- Gropper, George L. Does "Programmed" Television Need Active Responding? *AV Communication Review* 15: 5-22; 1967.
- Gropper, George L. Programming Visual Presentations for Procedural Learning. *AV Communication Review* 16: 33-56; 1968.
- Gropper, George L.; Lumsdaine, Arthur A.; and Shipman, Virginia. *Improvement of Televised Instruction Based on Student Responses to Achievement Tests*. Studies in Televised Instruction, Report No. 1. Pittsburgh, Pa.: Metropolitan Pittsburgh Educational Instruction Stations and American Institute for Research, 1961.
- Grotelueschen, Arden. Effects of Differentially Structured Introductory Materials and Learning Tasks on Learning and Transfer. *American Educational Research Journal* 5: 191-202; 1968.
- Hartley, James. Research Report. *New Education* 2: 29-41; 1966.
- Hartley, James and Woods, Patricia M. Learning Poetry Backwards. *National Society for Programmed Instruction Journal* 7: 9-13; 1968.
- Johnson, Rita B. The Effects of Prompting, Practice and Feedback in Programmed Videotape. *American Educational Research Journal* 5: 73-79; 1968.
- Kanner, Joseph H. *The Instructional Effectiveness of Color in Television: A Review of*



- the Evidence. Stanford, Calif.: ERIC Clearinghouse on Educational Media and Technology, Stanford Univ., 1968.
- Karraker, R. J. Knowledge of Results and Incorrect Recall of Plausible Multiple-Choice Alternatives. *Journal of Educational Psychology* 58: 11-14; 1967.
- Kopstein, Felix F. and Seidel, Robert J. Comment on Schurdak's "An Approach to the Use of Computers in the Instructional Process". *American Educational Research Journal* 4: 413-16; 1967.
- Kress, Gerald C. and Gropper, George L. Non-Adaptive Consequences of Self-Pacing in Programmed Instruction. *National Society for Programmed Instruction* 5: 14-15; 1966.
- La Giapa, John J. Programmed Instruction, Teacher Ability, and Subject Matter Difficulty. *The Journal of Psychology* 68: 257-260; 1968.
- Levonian, Edward. Retention of Information in Relation to Arousal During Continuously Presented Material. *American Educational Research Journal* 4: 103-16; 1967.
- Levonian, Edward. Auditory and Visual Retention in Relation to Arousal. *AV Communication Review* 16: 57-62; 1968.
- Mager, Robert F. *Developing Attitude Toward Learning*. Palo Alto, Calif.: Fearon, 1968.
- Markle, Susan M. Empirical Testing of Programs. *Programed Instruction: Sixty-Sixth Yearbook of the National Society for Study of Education*. Chicago: Univ. of Chicago Press, 1967. Chapter 5, pp. 104-138.
- Melching, William H. and Nelson, Frank B. *The Influence of Practice Frames and Verbal Ability on Programed Instruction Performance*. Technical Report 66-1. Alexandria, Va.: Human Resources Research Office (300 North Washington Street), George Washington Univ., 1966.
- Meyer, Donald E. *A Comparison of Response Confirmation Techniques for an Adjunctive Self-Study Program*. Aerospace Medical Research Laboratories. Report No. AMRL-TR-66-98 Springfield, Va.: Clearing House for Federal Scientific and Technical Information, 1966.
- Moore, James C. Manipulating the Effectiveness of a Self-Instructional Program. *Journal of Educational Psychology* 59: 315-19; 1968.
- Nagel, Thomas S. *Effects on Achievements and Attitudes of Two Writing Styles Used With Programmed Materials*. Paper presented to the Amer. Educ. Res. Assoc., 1968, in Chicago. East Lansing: Mich. State Univ. (Mimeo.)
- National Education Association, Research Division. *Instructional Resources in the Classroom*. NEA Research Bulletin 45: 75-77; 1967.
- Neidt, Charles O. Use of Videotaped Instructional Television for Teaching Study Skills in a University Setting. *AV Communication Review* 15: 269-84; 1967.
- Newmark, Gerald and Sweigert, Ray L., Jr. Three Media in Search of Their Message: A Summary of Research on Three Approaches to Teaching Spanish. *AV Communication Review* 15: 285-303; 1967.
- Niedermeyer, Fred C. The Relevance of Frame Sequence in Programmed Instruction: An Addition to the Dialogue. *AV Communication Review* 16: 301-17; 1968.
- Payne, David A.; Krathwohl, David R.; and Gordon, John. The Effect of Sequence on Programed Instruction. *American Educational Research Journal* 4: 125-32; 1967.
- Popham, W. James. Instructional Idea Tapes in Teacher Education. *AV Communication Review* 14: 371-77; 1966.
- Robeck, M. *A Study of the Revision Process in Programmed Instruction*. Master's Thesis. Los Angeles: Univ. of Calif., 1965.
- Rothkopf, Ernst Z. Learning from Written Instructive Materials: An Exploration of the Control of Inspection Behavior by Test-Like Events. *American Educational Research Journal* 3: 241-49; 1966.
- Sawiris, M. Y. *An Experimental Study of Individual and Group Learning Using a Linear Geometry Program*. *Programmed Learning and Educational Technology* 3: 146-53; 1966.
- Schurdak, John J. An Approach to the Use of Computers in the Instructional Process and an Evaluation. *American Educational Research Journal* 4: 59-73; 1967.
- Scriven, Michael. The Methodology of Evaluation. *Perspectives of Curriculum Evaluation*. (Edited by Robert E. Stake.) AERA Monograph Series on Curriculum Evaluation No. 1. Chicago: Rand McNally, 1967.
- Severin, Werner. The Effectiveness of Relevant Pictures in Multiple-Channel Communications. *AV Communication Review* 15: 386-401; 1967.



- Tiemann, Philip W. *Student Use of Behaviorally Stated Objectives to Augment Conventional and Programmed Revisions of Televised College Economics Lectures*. Paper presented to the Amer. Educ. Research Assoc., 1968, in Chicago. Chicago: Univ. of Ill. at Chicago Circle. (Mimeo.)
- Tiemann, Philip W.; Paden, Donald W.; and McIntyre, Charles J. *An Application of the Principles of Programmed Learning to a Televised Course in College Economics*. Washington, D.C.: U. S. Department of Health, Education, and Welfare, Office of Education, Title VII Project No. 5-0841, 1966.
- Tobias, Sigmund. *Lack of Knowledge and Fear of Automation as Factors in Teachers' Attitudes Toward Programmed Instruction and Other Media*. *AV Communication Review* 14: 99-109; 1966.
- Tobias, Sigmund. *Dimensions of Teachers' Attitudes Toward Instructional Media*. *American Educational Research Journal* 5: 91-98; 1968.
- Traub, Ross E. *Importance of Problem Heterogeneity to Programed Instruction*. *Journal of Educational Psychology* 57: 54-60; 1966.
- Travers, Robert M. W. *Research and Theory Related to Audiovisual Information Transmission*. Revised Edition, 1967. Kalamazoo: Distributed by Western's Bookstore, Western Michigan University.
- Unwin, D. and Leedham, J. (editors). *Aspects of Education and Technology: Proceedings of the Loughborough Programmed Learning Conference of April 1966*. London: Methuen, 1967.
- Vander Geer, A. W. and Thorne, Harvard E. *An Investigation of the Improvement of Informational Filmstrips and the Derivation of Principles Relating to the Effectiveness of These Media: Study No. I*. University Park: College of Education, Penn. State Univ., 1964a.
- Vander Geer, A. W. *An Investigation of the Improvement of Informational Filmstrips and the Derivation of Principles Related to the Improvement of These Media: Study No. II*. University Park: College of Education, Penn. State Univ., 1964b.
- Vander Geer, A. W. and Montgomery, R. E. *An Investigation of the Improvement of Informational Filmstrips and the Derivation of Principles Relating to the Improvement of These Media: Study No. III*. University Park: College of Education, Penn. State Univ., 1964c.
- Vander Geer, A. W.; Morrison, Jack; and Smith, Philip. *An Investigation of the Improvement of Educational Motion Pictures and a Derivation of Principles Relating to the Effectiveness of These Media*. University Park: College of Education, Penn. State Univ., 1965.
- Wendt, Paul. *Optional Returns from the By-Pass*. *National Society for Programmed Instruction Journal* 6: 4-5; 1967.
- Williams, Joanna P. *Combining Response Modes in Programed Instruction*. *Journal of Educational Psychology* 57: 215-19; 1966.
- Wodtke, Kenneth H. et al. *The Effects of Subject-Matter and Individual Difference Variables on Learning From Scrambled Versus Ordered Instructional Programs*. Paper presented to the Amer. Educ. Res. Assoc., 1968, in Chicago. University Park: Penn. State Univ.

## AUTHOR

- W. JAMES POPHAM *Address*: University of California at Los Angeles, California  
*Title*: Associate Professor of Education *Age*: 38 *Degrees*: A. B., Univ. of Portland; M.ED., Univ. of Portland; Ed.D., Indiana Univ. *Specialization*: Teaching Proficiency.

## 4: CURRICULUM EVALUATION

ROBERT L. BAKER\*

*Southwest Regional Laboratory for  
Educational Research and Development*

The past three years have netted little in the way of definitive empirical research dealing with curriculum. To be sure, many papers have been written on the evaluation of specific programs, and reports have been made of the differential efficacy of a variety of instructional procedures. However, curriculum research and evaluation as a subject matter has not yet been defined and bounded.

### *Curriculum Research*

#### *The Conceptual Domain*

Goodlad and Richter (1966) developed an elaborate conceptual system for dealing with problems of curriculum and instruction. They lamented that there is no carefully engineered framework for identifying relationships among complex and interacting phenomena. They expressed concern not for the development of theory per se, but rather for the development of a system that is more than a theory in scope but less than a theory in precision and prediction. In other words, curriculum must be differentiated into a definable and researchable subject matter. Goodlad and Richter proposed a five category conceptual system that suggests a broadly defined data base and includes categories related to both processes and products of curriculum. Although curriculum was defined as a set of intended learnings, the conceptual system includes levels of decision making in terms of remoteness from the learner, substantive curriculum decisions, type of decision to be effected at each level and between levels, identification of appropriate data sources to be consulted, and a clarification of authority and responsibility for decisions.

Bloom (1966) also supported the position that the construction of a clear map from which implications for research and evaluation may be drawn is a critical and necessary step. Although Bloom identified a few

\*Dr. Marvin D. Glock, Cornell University, served as the consultant to Dr. Baker on the preparation of this chapter. Dr. Baker is also indebted to Mmes. Allene Fishman and Judith Hostetler, Southwest Regional Laboratory for Educational Research and Development, who coordinated the initial survey for the research on which this chapter is based.

substantial advancements in curriculum research during the last 25 years, he suggested the need for establishing a better analysis system and more efficient networks for rapid communication of research discoveries.

*Additional References:* Guba (1967); Johnson (1967); Morriset and Stevens (1967).

### *Design Considerations*

If a technology for curriculum research and development is to evolve, researchers must be able to specify the subject matter with which they are dealing. In addition, the tools of the technology and the differential requirements for their use must be specified. Too often, the preoccupation with satisfying the requirements of design and statistical models violates the instructional treatment and reduces the utility of the research to zero. Conversely, the lack of specificity of treatment has often made the application of elegant procedures a waste of time and money at best and a smokescreen at worst. The requirements for both treatment specification and design must be considered carefully to ensure useful research results. Some gains in these areas have been made in the past few years. Novel research designs and measurements techniques suitable for employment in a variety of instructional situations have appeared in the literature (Campbell, 1965). The direct application of quasi-experimental designs to experimentation in educational settings was further elaborated by Stanley (1967). Stanley's major consideration was the lack of control in the typical educational setting. He cited four sources of probable internal experimental invalidity and at least one threat to the generalizability or external validity of such studies conducted in the typical educational setting. In proposing two designs of his own, Stanley affirmed that a researcher can compromise somewhat the requirements of true experimentation and still tease out reasonably unbiased causal data and conclusions.

Wiley (1967) supported the notion that educational researchers should substitute real experiments for the quasi-experiments suggested by Campbell, Stanley, and others. His reason is that quasi-experiments are laden with such gross problems of validity that reasonable inferences are unlikely to be drawn. Hall and Dzuiban (1967) pointed out that even when educational researchers are enjoined to use real experiments, the infeasibility of subjecting the studies to appropriate controls seldom permits such experimentation. They recognized that the compromises forced upon research design may yield spurious results, and pointed to a project of their own that was, as they termed it, "emasculated" by certain design flaws. Considerable doubt was cast on the validity of the resulting data. In this particular study, strong bias was operating because the experimental group had been self-selected rather than randomly chosen. Hall and Dzuiban recognized that this influenced the observation or the degree of internal validity; that external validity was academic; and that implications for

further study could be the only conclusion warranted.

Burnham (1968) pointed out that many official guidelines for research design and evaluation suggest designs which may lead to validity and reliability threats. He was particularly sensitive to the frailties of the proposed testing instruments and the presence of common sources of internal and external invalidity which remain generally uncontrolled by the design. Although his critique was too general to provide procedural cues, Burnham offered several courses of action to improve the validity of educational experiments: greater sophistication of methods for identifying, collecting, and analyzing data; reconsideration and formulation of the behaviors to be observed and the instruments to be used for observation; and a reconsideration and reformulation of the statistical analyses used in conjunction with the new designs and experiments. Burnham appended an unamplified suggestion that the sources of resistance to experimentation and evaluation by educational administrators must be dealt with.

Wardrop (1968) pleaded for greater emphasis on controlled experimentation in multi-classroom settings. He felt that although such experimentation would place greater demands on the resources of the school, it would help overcome some of the legitimate criticisms of contemporary educational research. The major limitation of current research cited by Wardrop is that statistical analyses have been based on individual students' performances when the experimental unit should have been defined as the classroom. If this is true, then the unit of sampling in educational research should be the class, the school, or the community, rather than the pupil. Although the examples Wardrop used to substantiate his point are more akin to evaluation procedures, they nonetheless helped to clarify the point. If one wishes to determine the effectiveness of a curriculum package or a teaching method administered by a teacher to a class, or of a scheduling system in which groups of students are assigned to either modular or traditional class schedules, one is faced with an experiment in which the experimental unit and hence the unit of analysis is a group or class of students and not the individual student.

Raths (1967) purported to support Wardrop's position in the description of a study in which the variance of scores of students who learned in a group was significantly smaller than the variance of scores of students who had learned individually. The logical support for Raths's position is unquestionable. However, since he was writing about performance variance resulting from individual and group learning situations, the empirical support appears to relate more to treatment specification than to the basis on which students are assigned to treatments. The point is that the experimental units should be the smallest units of students in the study to which treatments have been assigned randomly and which have responded independently of each other for the duration of the experiment. An additional discussion of the "analysis unit" problem may be found in Chapter 2 of this issue of the *Review*.



In an attempt to overcome several of the shortcomings of experimental arrangements typically found in an educational setting, Quilling (1968) developed a rather elaborate set of procedures. "Research and Instruction Units" were defined for schools in which teachers and pupils typically change rooms as pupils are regrouped for instruction throughout the day. Experimental arrangements requiring random assignment of pupils to groups in this situation served to minimize any reaction with the treatment as compared to the risk involved when pupils are assigned to self-contained classrooms. Where flexible scheduling and team teaching are the norm, such special experimental arrangements should have a minimum effect on the results. The "R and I Units" were organized so that true experimentation in a school setting became possible. The special staffing and physical arrangements may introduce another kind of invalidity; however, the promise for the future of doing "iterative" and "additive" or "subtractive" experimentation more than offsets the threats of error.

The problems of conducting experiments are not restricted to research on the conventional cognitive problems. Harrison (1967) reviewed the problems in design and interpretation of research on human relations training. The problems that he encountered and listed are almost identical to the problems reflected in other studies. Wodtke (1967) published a paper on the difficulties of interpreting retention data in educational experiments. He identified several established experimental designs that are appropriate for the study of retention effects. This seems to be particularly timely, since the notion has been frequently advanced that measures of retention taken immediately after acquisition are inadequate as criteria for educational improvement. In conjunction with the idea that it is necessary to determine the interdependency and interrelationship among various educational achievements, Wodtke made a justifiable suggestion that the effects of experimental treatments on retention over a longer period of time need greater emphasis in research.

Goodwin (1966), Dayton (1967), Fanning (1968), and Cook (1968) empirically specified some problems that arise from special conditions associated with an experiment. In these studies such extraneous variables as effects of experimental atmosphere, notice of testing, test administration and test scoring were isolated and their influence on the precision of the data collected after classroom experimentation was determined. Results from Goodwin's study showed that the "experimental atmosphere" treatment groups received higher scores, but the differences were not statistically significant. The "notice of testing" source of variation was the only one to attain significance, and this occurred on only one of the subtests. The obvious conclusion was that many extraneous variables related to experimental procedures had little or no effect on the dependent measures used in the study. Dayton (1967) and Fanning (1968) found that experimenter bias adversely influenced the validity of a great deal of research in education.

Cook (1968) reported the results of a four-year study of the Hawthorne effect. Conclusions reached at the end of an exhaustive literature search and an empirical study of elementary school children cast some doubt on the existence of the Hawthorne effect. Cook's search of the literature revealed that the existence of the Hawthorne effect in educational research is based more on intuition and logic than on solid evidence; his empirical study of the phenomenon among elementary school children yielded no evidence of the Hawthorne effect. Cook concluded that educational researchers would do well not to worry about the Hawthorne effect until evidence proves its existence and importance. Perhaps the Hawthorne effect has something in common with the legendary Yeti. If so, the researcher need not go in active search of the beast, but neither should he leave his camp unprotected as if the "Snowman" did not exist.

There is little doubt that most of the "extraneous" variables had little effect on student performance in the above referenced studies. There is probably little reason to expect otherwise. Even if the variables were potentially of a magnitude to make a difference, the error variance introduced by frail instruction and loosely controlled classroom delivery could easily account for the lion's share of the total variance. This is not an attack on the instruction proffered in the studies cited because no information on the instruction is available. However, a frail experimental vehicle will often doom a study to the production of no useful information despite the potential power of the variable under study.

### *Methodological Considerations*

Researchers appear to readily accept the legitimacy of the demand for more tightly controlled educational experiments in a natural environment and the recommended use of multiple-hypotheses procedures in experimental designs to permit the testing of several hypotheses simultaneously. Longitudinal studies involving careful analyses of retention, interaction, and sequencing effects have also been touted but ignored in practice. The creation of special purpose research and instruction units to permit adequate experimental control has been advanced as a partial solution to the reduction of uncertainty and error. Redefinition of the experimental unit of observation to insure a valid probability statement about the chances of drawing false conclusions from the data promotes greater confidence in the results. However, it magnifies the problem of securing sufficient experimental units and this is likely to reduce the probability that factorial designs will be widely used.

Optimism is reflected in papers containing direct attacks on problems of treatment specification and definition of dependent variables and measures. In an evaluation context, Wittrock (1966) shed light on useful approaches to the specification of instructional treatment. He emphasized the necessity for specifying precisely the instructional variables and avoiding error-filled complex instructional treatments. He expressed his preference

for getting at the complexity more systematically by studying interactions between instruction and the various contexts in which it is employed. His exemplary study involved the random introduction of two different mathematics curricula into several schools. The identification and assessment of the predetermined contextual variables in the various schools permitted an analysis of the independent variables (both instructional and contextual) and the dependent variables. This variation of a factorial model enabled Wittrock to assess the instructional and contextual variables differing across treatments.

Scandura (1966) reported an attempt to define a scientific set function language (SFL) that would aid the researcher in formulating research questions and in defining more precisely the dependent variable. This language was developed for research on mathematics teaching. The basic characteristic of SFL is that the principle rather than the association is defined as the basic behavior unit of knowledge. Knowledge in this context is that which makes possible the selection of the correct response in a class of responses to each stimulus in a class of stimuli. Although specifying precisely certain unit characteristics of behavior is somewhat easier in mathematics research than in other cognitive areas, such an attempt does have potential for more precise dependent variable definitions in other subject matter areas.

Staying within the multi-hypotheses factorial design frame, McLean (1966) presented some useful alternatives for effectively meeting such problems of research in a school setting as the presence of many dependent variables worthy of study, multiple sources of error, difficulty in obtaining experimental units, and miscellaneous administrative problems. Drawing on logical and empirical evidence, McLean concluded that fractional factorial designs can increase experimental efficiency. There are several approaches to the design of fractional factorial studies. Basically, such designs permit the experimenter to obtain information concerning main effects and certain interactions with a fraction of the cases required for a complete factorial design by sacrificing information on effects of variables and interactions of minimal interest.

Walberg and Welch (1967) described a fractional sampling technique developed for the quasi-experiment being conducted by the evaluation group of Harvard Project Physics. Although the technique is probably more useful for large scale experimental programs, it has application in experimental research for which multiple dependent variables are required and in which the measurement procedures are somewhat cumbersome. The Walberg-Welch plan is to draw a random sample of teachers from a national pool of volunteers in Project Physics. The data collection is conducted within intact classes drawn on a random basis. Several different tests are administered simultaneously to one class in which not all children take the same test.



An item sampling technique (Husek and Sirotnik, 1967) for obtaining experimental or evaluative data on class or group performance might prove especially effective with change or growth studies in which ordinarily it would be prohibitively expensive to take the measurement samples required. This technique provides a fractional sampling of people and items. The data from the sampling of people taking a sample of items is used to draw inferences about the performance of a population of subjects taking a population of items. The technique is potentially applicable when a large number of measures are required for a fairly small number of experimental or evaluation units.

Pruzek and Kleinke (1967) classified 298 education and psychological research papers, published in five journals during 1966-67, on the basis of the authors' purposes, designs, assumptions, and analyses. Their purpose was to determine whether the consistencies and inconsistencies in statistical models might influence the kinds of inference or conclusion drawn. In their opinion, the recent proliferation of statistical methodology has necessitated research on whether different methods lead to substantially different interpretations of data and on whether the algorithms used when machines and programs of different designs are employed lead to discrepant solutions. Their hunch is that there are likely to be discrepancies. A reasonable hunch not acknowledged by these authors is that if the instructional treatments specified are so feckless that the variance attributable to them is washed out by the idiosyncracies of machines and designs, curriculum research is indeed in trouble.

There is no question that the multi-criterion, multi-hypothesis, and complex experimental control requirements are going to demand new and dramatic approaches to educational experimentation. Buchanan (1967-68) introduced the idea that cybernetic concepts might be of potential value to educational research. He emphasized the application of cybernetic concepts to such an untouched research area as interpersonal communications. He also suggested that the complex requirements in educational research, particularly as they relate to considerations of methodology and possible implications for computer technology, will require such application.

The authors of the previous papers have posed some serious problems in the conduct of research in an educational setting; they also have presented some useful partial solutions to the persisting problems. However, conditions in the natural setting are likely to preclude any really serious attempts to conform more closely to "acceptable" procedures for experimentation. Guba (1965) recognized that there are questions to be asked and solutions to be generated in reality; he suggested strategies for conducting experimental observations or field studies when experimental or laboratory-type studies are not feasible or acceptable. He reported a number of field study techniques and added his suggestions for introducing the rigor that is possible or acceptable under such conditions. Eboch (1967) also saw the field study as a kind of middle ground between what he



called "mechanistic methodology" and "puritanical publicity." He recognized the need to conduct research that is closer to reality than many conventional research studies and closer to "objectivity" than the typical journal article. He wanted to promote field specialist "objectivity" without requiring hypotheses, controls, and pretests.

Judging by all that has preceded, one would guess that the educational researcher is hopelessly pinned beneath the sword of Damocles. There is little question that if one could eliminate all sources of invalidity and all experimental error, conclusions derived from educational experimentation would provide an infallible basis for educational practice. Since this is not possible, the best one can do is to attempt to minimize both and give them whatever consideration is possible in the statistical analyses; but the educational practitioner must be able to discriminate between the statistically significant and the practically significant result. An independent variable is practically significant only to the degree that it accounts for the variance of the dependent variable.

Schutz (1966) applied a statistical index termed "omega squared" (Hays, 1963) to all of the articles published in the *Journal of Educational Psychology* and the *American Educational Research Journal* in 1964 that reported a statistically significant  $F$  or  $t$  test. Omega squared reflects the predictive power afforded by a given functional relationship. When omega squared is 1.00, knowledge of the independent variable lets us know precisely the dependent variable. As omega squared approaches 0, knowledge of the independent variable provides less uncertainty reduction about the dependent variable. Schutz found the exercise a "disappointing display." It was not uncommon to find that the omega squared for a significant  $F$  value of, say, 22.0 was less than .15, indicating that less than 15 per cent of the total dependent variable variance was explained in the experiment. Although Glass and Hakstian (1969) questioned the mathematical base of omega squared, the statistic appears to provide a more accurate representation of the practical significance of the results of an experiment than the traditional  $F$  value. It is silly to make decisions involving thousands of people and millions of dollars on the basis of a highly significant  $F$  value only to find that less than 15 per cent of the total variance associated with the dependent variable has been explored.

## Curriculum Evaluation

### *Purpose and Scope*

The ambiguity that exists in the distinction between curriculum research and curriculum evaluation activities is likely to increase with the growing sophistication of researchers in dealing with "process" and "product" in a single study. Goodlad and Richter (1966) pointed to the importance of longitudinal studies and the application of computer tech-

nology to the study of highly complex and interrelated variables. They would approve almost anything to shake loose from the forced and traditional kind of experimentation that involves one independent variable, one dependent variable, a pretest, a posttest, and (occasionally) a delayed retest.

Although logically, and perhaps semantically, one can distinguish between research and evaluation, it is not very useful to base that distinction on a review of the literature dealing with curriculum research and evaluation. A distinction between curriculum research and curriculum evaluation on the basis of the terms used by the authors in their titles and the operational definitions assigned to the terms in the articles is nearly impossible.

The literature reflected fairly common agreement at a global level on the purposes of evaluation activities. Beyond that, however, there seemed to be considerable divergence on the subject matter of evaluation, its focus, and methodology (Hastings, 1966; Wilhelms, 1967; Bloom, 1967; Glaser, 1967; Scriven, 1967). Establishing an empirical basis for the revision and refinement of facilities, materials and methods appears to be a common objective of evaluation models.

In support of ambiguity, but certainly not in support of sloppy definitions of terms, Bloom (1967) suggested that in evaluation studies the environment (instruction, class, or school) is assumed to be a major source of the behavioral changes. Therefore, evaluation should be as much concerned with the characteristics of the environment which produce the changes in student behavior as it is with the appraisal of the changes themselves. This position certainly suggests that in evaluation activities considerable attention must be paid to specifying environmental conditions that interact in producing the payoff response of interest. In the first AERA monograph on curriculum evaluation, Gagné (1967a) defined curriculum in terms of student achievement, but he also described the importance of determining the relation of student achievement to the selection and organization of content. Aside from the implications for evaluation activities, Gagné's suggestions for research involved the design of curriculum before it has actually been used with students and a subsequent trial and revision cycling. Gagné also gave useful suggestions for controlling content in extended sequence learning research and evaluation.

Scriven (1966, 1967) warned against diluting evaluation activities by failing to distinguish between the goals and roles of evaluation. The specific role in evaluation may condition the form in which the goal activity takes place, but it should never deter efforts to gather and combine the performance data necessary to a data base for making decisions and judgments.

Wilhelms (1967, p. 15) posed five basic questions and condemned present evaluation activities for being inadequate.

... it [present evaluation activity] is equally to be condemned for the narrowness of its focus because in its gross exaggeration of the more mechanical, easier to measure features of education, it virtually blots the broader more fundamental objectives out of sight. The end result is not simply bad evaluation, it is distorted teaching and learning.

Wilhelms set forth two broad proposals for a new evaluation system. He proposed the continuing use of feedback to answer the larger questions of curriculum and policy. His other proposal dealt with the problems of instructional diagnosis that the teacher needs to sharpen his teaching and equally involves the active participation of the learner in the process. In the context of Scriven's "goal-role" distinction, the Wilhelms proposal creates a certain amount of confusion. Although there is no argument about the legitimacy of Wilhelms's proposal as a framework for defining evaluation activities, there is a confounding of values, goals, and roles of evaluation with the methodology of evaluation. Wilhelms would have evaluators playing one of the roles of evaluation instead of seeking its goal.

Wilhelms accurately perceived some inadequacies in extant evaluation procedures and identified some potentially useful objectives to build into the curriculum. Glaser (1967) did essentially the same thing, but he cast his procedures into an instructional model that provides for a direct attack on defining criterion categories and specifying instructional strategies. He made the defining statement that the current educational trend points to an increasing emphasis on cognitive development in the disciplines. That emphasis will be placed on developing continuity of education throughout the life-span of the individual. This will necessitate the development of the individual's ability to know how to learn and how to teach himself. Consequently, the onus will be on the educational enterprise to develop procedures for the adaptation of instruction to individual requirements. Analysis of such a defining statement suggested the need for a model that will systematically accommodate the unique requirements of both goal and role definitions. Glaser provided a seven category model that included behavioral objectives and conditions, entry behavior, diagnosis, alternatives keyed to diagnostic profiles, monitoring and assessment procedures, critical categories to be optimized (e.g., retention, transfer, etc.), and procedures for system improvement. Similar structures were suggested and tested by Baker and Schutz (1967), Gagné (1967a) and Schutz (1968). Although all of the structures mentioned are directly related to pupil achievement and behavioral properties, the categories may be useful in handling other "effects" or dependent variables.

*Additional References:* Alkin (1967); Stufflebeam (1966); Taylor (1967).

### *Educational Objectives*

Some of the general problems that confront evaluators and that



should stimulate increased impetus for greater rigor and a more systematic approach were described by Lortie (1967). He stated that consideration of the strengths and weaknesses of evaluation to date has been fairly irrelevant; a study of history has shown that educational issues were resolved by vote rather than by recourse to professional opinions based on empirical study. It has been said before that public education is a continuing constitutional convention; however, emphasis on consensus does nothing to advance techniques of evaluation. Lortie pointed out the evidence provided by "protracted engagements in particular learning structures" that have neither logical nor empirical foundation. The educational Establishment has no tradition of tough-minded empirical evaluation; instead it has stressed the merits of intuitive judgment and lore. The occasional flirtations with behavioral objectives suggest that a great deal more effort has been expended in verbalizing objectives than in operationalizing them. Lortie's final editorial comment seems reasonable; he said that the continued unchecked controversy over the conduct of educational evaluation will likely result in a loss of public confidence—unless ways are found to limit conflict and get on with the task. In his opinion, the verbal broadening of educational objectives will make no discernible impact until specific instructional practices are aligned with specific social goals. These problems are magnified by the pressures placed on the Establishment to expand alternatives rather than to reduce them. As other researchers have pointed out, achieving this alignment is going to require an examination of new evaluation methods. Tradition is an unsuitable guide for choosing among competing novelties.

The introduction of innovations creates special and demanding evaluation problems because one has to consider local circumstances, interaction effects among innovations, etc. A number of authors (Gagné, 1967a; Bloom, 1967; Glaser, 1967) have identified the need for opening new dimensions, defining more useful dependent variables and effects, and establishing systematic procedures for longitudinal evaluation. These considerations become critical issues when they affect long-term studies associated with product development research (Schutz, 1968; Baker, 1967), and when emphasis is placed on evaluating the educational product while it is in the development or formative stages (Scriven, 1967).

Although it is not a product of just the past three years, recent emphasis on specifying instructional objectives operationally or behaviorally (Gagné, 1967a; Stake, 1967; Glaser, 1967; Popham, 1968) as a means for defining criteria and specifying instruction has had an impact on the verbal behavior of educational researchers. There is recognition also that the objective alone is no magic solution to the problems confronting the practitioner, researcher or learner. Eva Baker (1967) contrasted the behavioral and non-behavioral effect that stated objectives have on pupil learning. She concluded that irrespective of the characteristics of the objectives, the teacher must first be given specific training on how to use



objectives in specifying instruction, and then motivated to use them.

Although they recognize the importance of defining curriculum by developing a sequenced set of instructional objectives, science curriculum researchers have experienced problems with a discrepancy between the statement of the objectives or the proposed assessment procedures and the perceived meaning and intention of the curriculum developer or teacher (Grobman, 1966; Walbesser, 1966; Walbesser and Carter, 1968). Each of these researchers suggested pre-evaluation strategies that might help increase the efficiency and contribution of behavioral objectives to curriculum development and evaluation.

Atkin (1968) supported the notion that there are "benefits to be derived from attempts to rationalize our decision-making process through the use of behaviorally stated objectives." However, he issued a warning that the use of behavioral objectives for planning may be suitable for the Department of Defense, educational technologists, and researchers, but not for the curriculum planner-theorist. He believed that rigid adherence to specifying the behavioral outcomes of all instructional activities would tend to decrease their educational relevance and eliminate many other "worthwhile" experiences from the curriculum. It seems that Atkin has not acknowledged the possibility of using the "behavioral objective" strategy as a development tool to help cultivate the unknown, to define worthwhile but as yet "phantom" objectives, and to specify procedures for maximizing instructional relevance of objectives.

In response to the urgent need for criterion referenced instructional objectives, Popham (1968) and Popham and Skager (1968) developed the interesting concept of a depository of objectives, which eventually should promote more precise curriculum research and development. Recognizing the reluctance and inability of the consumer to specify his own objectives in sufficiently precise terms to accommodate valid criterion test development and assuming that the presence of already developed objectives would be appealing to the consumer, Popham and Skager are developing a national depository for instructional objectives and related measurement devices. At the outset, the primary service of the depository and exchange will be to make available alternative criterion referenced objectives and measuring devices to assist school personnel in instructional and evaluation activities. Popham and Skager's rather ambitious plans involve collecting, storing, processing, and distributing objectives and evaluation measures. One requirement is that test items alone will not be available as a selection option; the user will be required to select objectives. However, user preferences for various levels of item specificity and user ratings on the relevance of items under each condition of specificity will be solicited.

In the depository, a refined and continually updated basis for curriculum and evaluation will be established. A less significant implication relates to the potential large-scale assessment of curriculum activities among the users of the system. However, the large-scale assessment that might

result from this program will be of a different magnitude than that planned for the national assessment program (Tyler, 1966; Merwin and Tyler, 1966; Tyler, 1967). In this program, the objectives for assessment were formulated by a national panel of teachers, curriculum workers, and subject matter specialists. After they are reviewed by various lay panels from four geographic regions, the resulting assessment instruments will be used to appraise the educational progress of large populations. The feedback from this program is of a considerably different nature than the feedback from the Popham-Skager system.

*Additional References:* Alkin (1968); Leton (1966); Metfessel and Michael (1967a,b).

### *Instructional Specifications*

A behavioral objective and a criterion referenced test constitute the beginning and end of an instructional segment. For the development of a sensitive evaluation plan, analyses of the distinctions among instructional antecedents, transactions and outcomes must be made (Stake, 1967). In one of the early efforts to provide a standardized set of procedural cues, Baker and Schutz (1967) employed a system for guiding activities in the analysis of the behavioral properties of an instructional objective and further specifying, ordering, and sequencing instruction within the bounds of the objective. Their simple paradigm is an extension and adaptation of the work done by Hively (1963) on item forms and is consistent with the work of Glaser (1967) and Gagné (1967b). The Schutz-Baker framework provided cues for the construction of psychometrically adequate criterion instruments, mastery tests, and revision and sequencing strategies for each behavioral objective. Four specific categories were analyzed for each objective:

- (1) Instructional cues required—the principles, rules, and concepts to be developed as a means for verbally mediating the desired response;
- (2) Elicitors specified—prototype items to be used in sampling the desired behavior;
- (3) Stimulus limits defined—delimiting and distinguishing between stimulus conditions in which the desired response would be appropriate and stimulus conditions in which the desired response might occur but would be inappropriate;
- (4) Entry behavior—behaviors and skills the learner should possess prior to entering a phase of instruction.

The procedures proved to be quite useful in specifying instruction and also aided in the systematic manipulation of instructional conditions and procedures.

The implications for the development of criterion instruments are most crucial at the moment. The emphasis on mastery and the reflection of

student performance against an absolute criterion make many of the conventional psychometric procedures inappropriate (Ebel, 1966; Tyler, in Stake, 1967; Husek, 1968; Stake, 1968). Pitfalls in the use of conventional or standardized tests as they relate to content validity have already been mentioned. Because current tests seek to measure individual differences among pupils taking the test, the tests concentrate on items which differentiate the children. Very difficult and very easy items are eliminated because they do not afford sufficient discrimination. Conventional norm reference tests are designed to maximize the variability among those tested; criterion referenced tests are designed to minimize inter-subject variance and to focus on the specified learned behaviors of the students tested. Associated problems related to item construction, the reliability and validity concepts and interpretation of data from criterion referenced tests, need considerable attention from psychometric specialists. Tyler (1967) stated that success in school can no longer be used as an acceptable criterion for validating predictive measurement instruments. He elaborated on the statement by suggesting that the general objectives and aims of education have changed to the extent that by definition all children will succeed. The changing structure of the labor force and higher requirements for intelligent citizenship demand this success.

A clarified semantic distinction between the outcomes or products of instruction and the conditions of instruction will be of great benefit to the research and evaluation enterprises. Although the labels attached to both vary considerably, their referents have much in common. Fairly clear distinctions have been made between procedures and accomplishments, processes and outcomes, conditions and behaviors, means and ends, and transactions and outcomes. Common to all is the notion that although researchers are vitally interested in determining a set of procedures for appraising changes in students that have been produced by educational means, research methodologies must reflect clearly defined distinctions between means and ends. Bloom (1967) cited the search for evidence on both the individual and environment as the major strength of evaluation and assessment. The immediate necessity for studying various and complex interactions between process and product and the emerging computer capability for storing and manipulating enormous amounts of data make this requirement critical and feasible.

Glaser (1967), in distinguishing between the evaluation of procedures and the evaluation of accomplishments, and Scriven (1967), in defining the difference between intrinsic evaluation and payoff evaluation both emphasize the importance of making a clear distinction between and a systematic study of process and product. As the distinction becomes more clearly defined methodologically, attempts are being made to develop more precise specifications and requirements for the systematic study of environmental and instructional conditions as they relate to and interact with



student performance. For a number of years Pace (1968) has been concerned with the range of outcomes of higher education that may be observed in college-educated adults and with the nature of the educational experience itself. In his series of studies, he emphasized the development of contextual measures to be used in interactive studies with certain identified criterion measures. To date, he has developed three contextual measures: one related to campus morale, another related to the quality of teaching and faculty-student relationships, and a third on the analysis of student logs to see how students spend their academic time. These procedures have potential application for studies dealing with the interaction effects of institutional or environmental characteristics and specified conditions of instruction on student performance. Wittrock and Jones (1968), working directly in the classroom environment, specified procedures for the evaluation of classroom interaction and subsequent interactions with instructional conditions as effects on student performance. To date, Wittrock and Jones have directed their effort to the development of instruments that will produce usable indexes utilizing brief samples of classroom behavior. Reflection on the several papers reviewed above suggests that useful definitions and techniques are emerging which should help to separate experimentally the effects of pervasive environmental characteristics and conditions of instruction as they interact to influence student performance over a more extended period of real time learning.

### *Formative and Summative Evaluation*

Reacting to the tremendous variation in roles that evaluation activities play, a potentially useful distinction between two kinds of evaluation was developed by Scriven and summarized by Ahmann (1967). The qualifying labels, "formative" and "summative" were attached to the two evaluative procedures. Formative evaluation techniques are employed when one is interested in revising the curriculum while it is still in development. This means that evaluation activities must take place at predetermined stages in the development of the curriculum and that strategies must be included in those activities to permit changes to be made on the basis of reliable and valid criterion referenced evidence. This permits adjustments to be made in the curriculum on the basis of deficiencies identified enroute. It also suggests a trial-revision cycle based on predetermined standards of enroute student performance or a defined non-student outcome.

Summative evaluation techniques produce some kind of terminal or overall estimate. Ordinarily the purpose is to arrive at a general conclusion about a curriculum or instrument; one example is the determination of the relative merits of one textbook or course of study over another. The distinction between the two forms of evaluation is not that one is concerned with normative data and the other with criterion referenced data; indeed, both may employ an absolute standard of performance in evaluation. Perhaps a useful distinction cannot be made until it is determined who is going to



use the evaluation data and for what purposes. Ahmann (1967) also found the differentiation between summative and formative evaluation somewhat less clear than it appears to be.

Dick (1968) found formative evaluation efforts in practice to be difficult, but useful. In an empirical study using a programed calculus text he found that some procedural strategies need to be further refined before formative evaluation efforts will become a very powerful tool. It seems that a technology for collecting, analyzing, and interpreting data to writers, text specialists, and subject matter experts will have to be worked out before the materials revision improvements will become commonplace results of formative evaluation procedures. Sullivan (1969) emphasized the importance of precise instructional objectives and specifications, detailed records, and an elaborate management system in the development of formative evaluation procedures. In his "case study" account of developing an instructional product, he illustrated the necessity of effecting tight controls and careful monitoring of product development and revision activities.

"At present, formative evaluation methods have much the same status as the invisible needle and thread used by the tailors in the 'Emperor's New Clothes'" (Schutz, 1968). Pointing out that formative evaluation efforts have been highly content oriented, Schutz advocated shifting to a product orientation in which the criteria for gauging instructional improvement become straightforward and the application of formative evaluation procedures becomes more meaningful. He presented a four dimensional description of current school practice: print mediated, group paced, managed against relative norms, and teacher-based. Granting the desirability of shifting to multi-mediated, individually paced, managed by objectives, and computer based instructional practices, Schutz did not feel that the magic of words will effect the shift. Instead, he described a "primitive" model for product development research. This model may be described as two approaches involving a convergent methodology. The first is a trial revision cycle working within the boundaries of the present instructional system to create a product that will achieve predetermined objectives now! The second approach is a follow-up and introduces modifications in one or more of the major dimensions of practice, e.g., use of a computer based instructional management system. This latter is called a new "generation" of the product.

The emergence of a technology, although primitive, for the conduct of research in product development should serve to reduce the ambiguity of the relationship between curriculum research and evaluation activities. Preparation of materials and methods that will reliably effect student achievement of predetermined educational outcomes requires the systematic manipulation of instructional variables and conditions as well as continuous evaluation of their effects. Existing ambiguity must be replaced by the specification of the differential research and evaluation functions to be performed based on defined curriculum decision requirements.

*Additional References:* Berlak et al. (1967); Lindvall (1966); Taylor (1967).

### Bibliography

- Ahmann, J. Stanley. Aspects of Curriculum Evaluation: A Synopsis. *Perspectives of Curriculum Evaluation*. (Edited by Robert E. Stake.) AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally, 1967. Pp. 39-83.
- Alkin, Marvin C. *Evaluating the Cost Effectiveness of Instructional Programs*. From the Proceedings of the Symposium on Problems in the Evaluation of Instruction, Dec., 1967. CSEIP Occasional Report No. 25. Los Angeles: Univ. of Calif., Center for the Study of Instructional Programs. (Offset.)
- Atkin, J. Myron. Behavioral Objectives in Curriculum Design: A Cautionary Note. *Science Teacher* 35: 27-30; May, 1968.
- Baker, Eva Lee. *The Differential Effect of Behavioral and Nonbehavioral Objectives Given to Teachers on the Achievement of Their Students*. Doctor's thesis. Los Angeles: Univ. of Calif., 1967.
- Baker, Robert L. Research, Development and Dissemination Strategies in Improving Education. *Planning and Effecting Needed Changes in Education*. (Edited by Edgar L. Morphet.) Denver, Colo.: Designing Education for the Future: An Eight State Project, 1967. Pp. 106-115.
- Baker, Robert L. and Schutz, Richard E. *Technical Documentary: Research on Procedures for the Revision of Instructional Materials*. Wright-Patterson AFB, Ohio: Aerospace Medical Division, Air Force Systems Command, 1967.
- Bloom, Benjamin. Twenty-Five Years of Educational Research. *American Educational Research Journal* 3: 211-21; 1966.
- Bloom, Benjamin S. *Toward a Theory of Testing Which Includes Measurement-Evaluation-Assessment*. From the Proceedings of the Symposium on Problems in the Evaluation of Instruction, Dec., 1967. CSEIP Occasional Report No. 9. Los Angeles: Univ. of Calif., Center for the Study of Instructional Programs. (Offset.)
- Buchanan, Bruce H. Cybernetics and Educational Research. *Ontario Journal of Educational Research* 10: 73-81; Winter, 1967-68.
- Burnham, Brian. Facing Up to a Curriculum Research Design Problem: Evaluating Programs for the Educationally Deprived. *Ontario Journal of Educational Research* 10: 83-99; Winter, 1967-68.
- Campbell, Donald T. *Exploration of Novel Research Designs and Measurement Techniques*, 1965. (ERIC: ED 003 389.)
- Cook, Desmond L. Hawthorne Effect and Educational Research. *School and Society* 96: 40; January 20, 1968.
- Dayton, C. Mitchell. Implications of Educational Research of the Phenomenon of Experimenter Bias. *Educational Leadership* 24: 733-39; 1967.
- Dick, Walter. A Methodology for the Formative Evaluation of Instructional Materials. *Journal of Educational Measurement* 5: 99-102; 1968.
- Ebel, Robert L. Some Measurement Problems in a National Assessment of Educational Progress. *Journal of Educational Measurement* 3: 11-17; 1966.
- Eboch, Sidney C. The Value of Field Studies in Education. *Theory into Practice* 6: 69-72; 1967.
- Fanning, John F. Implications of Overt Manifestations of Expectancy Bias. *Educational Leadership* 25: 683-87; 1968.
- Gagné, Robert M. Curriculum Research and the Promotion of Learning. *Perspectives of Curriculum Evaluation*. (Edited by Robert E. Stake.) AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally, 1967a. Pp. 19-38.
- Gagné, Robert M. *Instructional Variables and Learning Outcomes*. From the Proceedings of the Symposium on Problems in the Evaluation of Instruction, Dec., 1967. CSEIP Occasional Report No. 16. Los Angeles: Univ. of Calif., Center for the Study of Instructional Programs. (Offset.)
- Glaser, Robert. *Theory of Evaluation of Instruction: Changes and Trends*. From the Proceedings of the Symposium on Problems in the Evaluation of Instruction, Dec., 1967. CSEIP Occasional Report No. 13. Los Angeles: Univ. of Calif., Center for the Study of Instructional Programs. (Offset.)
- Glass, Gene V and Hakstian, A. Ralph. Measures of Association in Comparative Ex-

- periments: Their Development and Interpretation. *American Educational Research Journal* 6: in press.
- Goodlad, John I. and Richter, Maurice N., Jr. *The Development of a Conceptual System for Dealing with Problems of Curriculum and Instruction*. Cooperative Research Project No. 454. Los Angeles: Univ. of Calif., School of Education, 1966.
- Goodwin, William L. Effect of Selected Methodological Conditions on Dependent Measures Taken After Classroom Experimentation. *Journal of Educational Psychology* 57: 350-58; 1968.
- Grobman, Hulda. The Place of Evaluation in the Biological Sciences Curriculum Study. *Journal of Educational Measurement* 3: 205-212; 1966.
- Guba, Egon. The Expanding Concept of Research. *Theory into Practice* 6: 57-65; April, 1967.
- Guba, Egon G. *Methodological Strategies for Educational Change*. Paper presented to conference on Strategies for Educational Change, November 8-10, 1965, in Washington, D.C. Bloomington: Ind. Univ., School of Education.
- Hall, Harry and Dzuiban, Charles. A Problem of Validity in Curriculum Research. *Educational Leadership* 25: 81-82; 1967.
- Harrison, Roger. *Problems in the Design and Interpretation of Research on Human Relations Training*. Paper presented to research seminar at Veterans' Administration Hospital, January 27, 1967, at West Haven, Conn. (ERIC: ED 014 015.)
- Hastings, J. Thomas. Curriculum Evaluation: The Why of the Outcomes. *Journal of Educational Measurement* 3:27-32. 1966.
- Hays, W. *Statistics for Psychologists*. New York: Holt, Rinehart, and Winston, 1963.
- Hively, W. *Defining Criterion Behavior for Programmed Instruction in Elementary Mathematics*. Cambridge, Mass.: Committee on Programmed Instruction, Harvard Univ., 1963.
- Husek, T. R. and Strottnik, Ken. *Item Sampling in Educational Research*, February, 1967. (ERIC: ED 013975.)
- Husek, Theodore R. Different Purposes for Tests and Their Implications for Test Construction and Validation. *Progress in Evaluation Study*. Third Annual Report to the U. S. Office of Education, Center for the Study of Evaluation. Los Angeles: Univ. of Calif., Graduate School of Education, 1968. Pp. 106-109.
- Lortie, Dan C. *The Cracked Cake of Educational Custom and Emerging Issues in Evaluation*. From the Proceedings of the Symposium on Problems in the Evaluation of Instruction, Dec., 1967. CSEIP Occasional Report No. 19. Los Angeles: Univ. of Calif., Center for the Study of Instructional Programs. (Offset.)
- McLean, Leslie D. Phantom Classrooms. *School Review* 74: 139-49; 1966.
- Merwin, Jack C. and Tyler, Ralph. What the Assessment of Education Will Ask. *Nation's Schools* 78: 77-79; November, 1966.
- Messick, Samuel. *The Criterion Problem in the Evaluation of Instruction: Assessing Possible, Not Just Probable Intended Outcomes*. From the Proceedings of the Symposium on Problems in the Evaluation of Instruction, Dec., 1967. CSEIP Occasional Report No. 22. Los Angeles: Univ. of Calif., Center for the Study of Instructional Programs. (Offset.)
- Morrisett, Irving and Stevens, W. Williams, Jr. Curriculum Analysis. *Social Education* 31: 483-86; 1967.
- Okay, James R.; Dick, Walter; and Westmeyer, Paul H. *The Use of Hierarchical Task Analysis in Formative Curriculum Evaluation*. Los Angeles: Univ. of Calif., Center for the Study of Instructional Programs. (Offset.)
- Pace, C. Robert. Evaluation in Higher Education. *Progress in Evaluation Study*. Third Annual Report to the U. S. Office of Education, Center for the Study of Evaluation. Los Angeles: Graduate School of Education, Univ. of Calif., 1968. Pp. 88-93.
- Popham, W. James. Instructional Objectives Exchange. *Progress in Evaluation Study*. Third Annual Report to the U. S. Office of Education, Center for the Study of Evaluation. Los Angeles: Graduate School of Education, Univ. of Calif., 1968. Pp. 110-112.
- Popham, W. James and Skager, Rodney W. Instructional Objectives Measurement System. *Progress in Evaluation Study*. Third Annual Report to the U.S. Office of Education, Center for the Study of Evaluation. Los Angeles: Graduate School of Education, Univ. of Calif., 1968. Pp. 113-115.



- Pruzek, Robert M. and Kleinke, David I. *Recent Developments in Educational Research Methodology*. Paper read at the 1967 Research Convocation of the Educational Research Association of New York, Nov., 1967 at Albany. Albany: State Univ. of N. Y., School of Education.
- Quilling, Mary R. *Controlled Experimentation in Research and Instruction Units. Research and Development Strategies in Theory Refinement and Educational Improvement*. Theoretical Paper No. 15. Madison: Univ. of Wis., Wis. Research and Development Center for Cognitive Learning, 1968. Pp. 10-13.
- Raths, James. The Appropriate Experimental Unit. *Educational Leadership* 25: 263-66; 1967.
- Scandura, Joseph M. Precision in Research on Mathematics Learning: The Emerging Field of Psycho-Mathematics. *Journal of Research in Science Teaching* 4: 253-74; 1966.
- Schutz, Richard E. The Control of "Error" in Educational Experimentation. *The School Review* 74: 150-58; 1966.
- Schutz, Richard E. Experimentation Relating to Formative Evaluation. *Research and Development Strategies in Theory Refinement and Educational Improvement*. Theoretical Paper No. 15. Madison: Univ. of Wis., Wis. Research and Development Center for Cognitive Learning, 1968. Pp. 19-22.
- Scriven, Michael. The Methodology of Evaluation. *Perspectives of Curriculum Evaluation*. (Edited by Robert E. Stake.) AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally, 1967. Pp. 39-83.
- Scriven, Michael. *The Methodology of Evaluation*. Social Science Education Consortium Publication 110, March, 1966. (ERIC: ED 014001.)
- Stake, Robert E. Countenance of Educational Evaluation. *Teachers College Record* 68: 523-40; 1967.
- Stake, Robert E. Testing in the Evaluation of Curriculum Development. *Review of Educational Research* 38: 77-84; 1968.
- Stanley, Julian C. Quasi-Experimentation in Educational Settings. *School Review* 75: 343-52, 367-68; 1967.
- Sullivan, H. J. *Instructional Objectives and Improved Learner Achievement*. Special Report 10. Inglewood, Calif.: Southwest Regional Laboratory for Educational Research and Development, 1969.
- Trow, Martin. *Methodological Problems in the Evaluation of Innovation*. From the Proceedings of the Symposium on Problems in the Evaluation of Instruction, Dec., 1967. CSEIP Occasional Report No. 31. Los Angeles: Univ. of Calif., Center for the Study of Instructional Programs. (Offset.)
- Tyler, Ralph W. Objectives for a National Assessment of Education. *Journal of Educational Measurement* 3: 1-4; 1966.
- Tyler, Ralph W. Changing Concepts of Educational Evaluation. *Perspectives of Curriculum Evaluation*. (Edited by Robert E. Stake.) AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally, 1967. Pp. 19-18.
- Walberg, Herbert J. and Welch, Wayne W. A New Use of Randomization in Experimental Curriculum Evaluation. *School Review* 75: 369-77; 1967.
- Walbesser, Henry H. Science Curriculum Evaluation: Observation on a Position. *Science Teacher* 33: 34-9; February, 1966.
- Walbesser, Henry H. and Carter, Heather. Some Methodological Considerations of Curriculum Evaluation Research. *Educational Leadership* 26: 53-64; October, 1968.
- Wardrop, James L. *Controlled Experimentation in Multiclassroom Settings. Research and Development Strategies in Theory Refinement and Educational Improvement*, Theoretical Paper No. 15. Madison: Wis. Research and Development Center for Cognitive Learning, Univ. of Wis., 1968. Pp. 6-9.
- Wiley, David E. *The Design and Analysis of Evaluation Studies: Comments and Suggestions*. From the Proceedings of the Symposium on Problems in the Evaluation of Instruction, Dec., 1967. CSEIP Occasional Report No. 28. Los Angeles: Univ. of Calif., Center for the Study of Instructional Programs. (Offset.)
- Wilhelms, Fred T. Evaluation as Feedback. *Evaluation as Feedback and Guide*. Washington, D.C.: Association for Supervision and Curriculum Development, 1967. Pp. 2-17.
- Wittrock, M. C. *The Experiment in Research on Evaluation of Instruction*. Dec. 1966. (ERIC: ED 012 107.)



- Wittrock, M. C. and Jones, M. H. The Evaluation of Classroom Interaction. *Progress in Evaluation Study*. Third Annual Report to the U.S. Office of Education, Center for the Study of Evaluation. Los Angeles: Univ. of Calif., Graduate School of Education, 1968. Pp. 76-80.
- Wodtke, Kenneth H. On the Assessment of Retention Effects in Educational Experiments. *Journal of Experimental Education* 35: 28-36. 1967.

### Additional References

- Alkin, Marvin C. *Towards an Evaluation Model . . . A Systems Approach*, August, 1967. (ERIC: ED 014150.)
- Alkin, Marvin C. A Simulated Evaluation Exercise. *Progress in Evaluation Study*. Third Annual Report to the U.S. Office of Education, Center for the Study of Evaluation. Los Angeles: Univ. of Calif., Graduate School of Education, 1968. Pp. 94-98.
- Berlak, Harold et al. *The Development of a Model for the Metropolitan St. Louis Social Studies Center*, Feb. 1967. (ERIC: ED 012 390.)
- Guba, Egon. The Expanding Concept of Research. *Theory into Practice* 6: 57-65; 1967
- Johnson, Mauritz, Jr. Definitions and Models in Curriculum Theory. *Educational Theory* 17: 127-40; 1967.
- Leton, Donald. *Criterion Problems and Curriculum Evaluation*. Paper delivered at the National Seminar for Research in Vocational Education, at Urbana, Illinois, May 16-20, 1966. (ERIC: ED 016 790.)
- Lindvall, C. M. The Task of Evaluation in Curriculum Development Projects: A Rationale and Case Study. *School Review* 74: 159-67; 1966.
- Metfessel, Newton and Michael, William. A Paradigm for Developing Valid Measurable Objectives in the Evaluation of Educational Programs in College and Universities. *Educational and Psychological Measurement* 27: 373-83; 1967a.
- Metfessel, Newton and Michael, William. A Paradigm Involving Multiple Criterion Measures for the Evaluation of the Effectiveness of School Programs. *Educational and Psychological Measurement* 27: 931-43; 1967b.
- Morrisett, Irving and Stevens, W. Williams, Jr. Curriculum Analysis. *Social Education* 31: 483-86; 1967.
- Stufflebeam, Daniel. A Depth Study of the Evaluation Requirement. *Theory into Practice* 5: 121-33; 1966.
- Taylor, P. A. A Theoretical Evaluation Model. *Educational and Psychological Measurement* 27: 305-21; 1967.
- Taylor, Peter A. The Use of Factor Models in Curriculum Evaluation: A Mathematical Model Relating Two Factor Structures. *Educational and Psychological Measurement* 27: 305-21; 1967.

### AUTHOR

- ROBERT L. BAKER Address: Southwest Regional Laboratory for Educational Research and Development Title: Director, Product Integration Age: 41 Degrees: B.A., Univ. of Neb.; M.A., Univ. of Neb.; Ph.D., Univ. of Neb. Specialization: Product Development, Instructional Strategies, Classroom Learning.

## 5: METHODOLOGICAL ISSUES IN CURRICULUM RESEARCH

RICHARD E. SCHUTZ

*Southwest Regional Laboratory  
for Educational Research and  
Development*

The preceding chapters reveal scant attention to methodological issues by curriculum researchers. The modal method of spirited typing does not readily generate issues of a methodological nature. Curriculum reformers and builders have tended to be amethodological as well as ahistorical.

Since methodological issues in curriculum research have not overtly appeared in the literature, this review will be structured in terms of latent issues of methodology currently available to the curriculum researcher.

### *Mythology or Reality?*

Curriculum developers have tended to assume an isomorphism between their interpretation of curriculum guides and instructional materials and that of teachers and students. As long as their assumption was not questioned, a gentlemen's agreement held. However, word is leaking out that a number of cherished curriculum beliefs are figments of the researcher's theoretical imagination. Nongraded classroom appear to be indistinguishable from graded classrooms (Goodlad, 1968); self-instruction in schools appears to be indistinguishable from teacher-instruction in schools, and endemic gaps are being identified "between emotionally-toned accounts of the ideology and the day-to-day reality in the life of innovative schools and classrooms" (Smith and Keith, 1968, p. 5). An overarching methodological issue is the invention of means to separate credibility from dependability. Gap-creating work appears to be necessary for separating advertising from data; and gap-filling work seems essential for separating conjectural from functional relationships involving educational phenomena.

### *Strong or Weak Inference?*

Platt (1964) observed that areas of scientific endeavor vary in productivity. His explanation (p. 347) for the high productivity of a few sciences is that "rapidly moving fields are fields where a particular method of doing

scientific research is systematically used and taught, an accumulative method of inductive inference" that Platt terms "strong inference." Strong inference is the systematic use of alternative explanations of available data, the performance of successive experiments to *exclude* as many of the alternatives as possible, and then the repetition of the process at the next fork in the conditional inductive tree. The regular and explicit use of alternative data-based hypotheses and sharp exclusions generates "logical trees" in which each new result provides a branch subject to alternative explanations and suggests experiments and controls that can reduce the number of explanations.

... the systematic use of strong inference suddenly gives us a yardstick for thinking about the effectiveness of scientific methods in general. Surveys, taxonomy, design of equipment, systematic measurements and tables, theoretical computations—all have their proper and honored place, provided they are parts of a chain of precise induction of how nature works. Unfortunately, all too often they become ends in themselves, mere time-serving from the point of view of real scientific advance, a hypertrophied methodology that justifies itself as a lore of respectability (Platt, 1964, p. 351).

There is little indication that curriculum research has made extensive use of strong inference. Many curriculum research studies appear to be included in the "sick science" category, yielding (Platt, 1964, p. 350) "simply a verbal formula which the graduate student repeats and believes because the professor has said it so often. This is not science but faith; not theory but theology."

Curriculum researchers have danced to the tune played by psychological research methodology. The lyrics of psychological research methodology have in turn been drawn from statistical analyses growing out of agronomy and biology. The disparity among the characteristics of agronomy, biology, psychology, and education was clearly noted by Glass (1968). Psychologists (for example, Sidman, 1960; Bakan, 1967) have occasionally penetrated methodological boundaries that the discipline in general regarded as sacred, but educational researchers have tended to join in efforts to plug the holes rather than in efforts to enlarge them (Schutz and Baker, 1968).

If one is interested in obtaining a comprehensive view of the fine-grain issues receiving current methodological attention, he can look over the shoulders of persons involved in conference debate on such matters. The Seventh Annual Phi Delta Kappa Symposium on Improving Experimental Design and Analysis (Stanley, 1966), the UCLA R&D Symposium on Problems in the Evaluation of Instruction (Wittrock and Wiley, in press), and the Educational Testing Service Invitational Testing Conference (1968) provide three rings for interested spectators to observe.

An antidote for the constriction of thought and vision created by the emulation of the psychologist-agronomist-biologist is provided by "general systems theory." To some the cure may be more painful than the constriction since by definition, if not design, systems thinking skirts intellectual unmanageability. A reasonable introduction is provided by Simon's (1962) analysis of complex systems, which are "made up of a large number of parts which interact in non-simple ways." Curriculum would unquestionably qualify as being within this definitional boundary. Indeed, a system is anything that is not chaos. Simon pointed out that "in the face of complexity an in-principle reductionist may be at the same time a pragmatic holist" and presented a conceptual-procedural framework for maintaining such a reductionist-holist stance, short of schizophrenia. The key to this trick is to analytically decompose complex systems into hierarchical subsystems. Each subsystem has integrity per se but also is an interdependent component of the system. Subsystems in turn have their own subsystems, and so on. More elaborate state-of-the-art presentations of systems theory and application are to be found in Buckley (1968) and Berrien (1968).

A didactic treatment of systems analysis was presented by Churchman (1968) who forthrightly acknowledged that "a great deal of nonsense has been written about the systems approach, because once an idea becomes popular it can be sold, and naturally some of its sellers are out to make a profit." His depiction of the systems approach via a debate between advocates and critics is ingeniously persuasive and instructive.

A cogent broad-band system analysis of education was conducted by Coombs (1968). His frame of reference is international, but his analytic framework can be extended by analogy to more localized educational systems.

### *Science or Technology?*

Curriculum researchers have expected their efforts to result in direct improvement in educational practice despite the fact that research reports typically conclude with suggestions for further research rather than with suggestions for implementation. When demonstrable educational improvement has not materialized after a relatively short interval, the popular ploy has been to blame institutional resistance to change or lore-based intellectual and personality deficiencies of educators. Much evidence suggests that this reasoning is unproductive and unrealistic.

First, it appears questionable that educational researchers have yet isolated variables that practitioners can manipulate to obtain predictable results. Stephens (1967) wrote a highly provocative summary of the summaries of experiments on instruction over the last fifty years. Stephens (1967, p. 9) commented on "the remarkable constancy of educational results in the face of widely differing deliberate approaches" and continued bluntly, "Every so often we adopt new approaches or new methodologies



and place our reliance on new panaceas. . . . Yet the academic growth within the classroom continues at about the same rate, stubbornly refusing to cooperate with the bright new dicta emanating from the conference room."

Replication of Stephens's spectacularly elegant non-significant difference and its generalization to the college level was reported by Dubin and Taveggia (1968, p. 45): "... we have reported the results of a careful and systematic reanalysis of the data of almost 100 comparative studies of different college teaching methods. We have found no shred of evidence to indicate any basis for preferring one teaching method over another as measured by the performance of students on course examinations." Education appears to be in a pre-scientific state rather than in a state of scientific revolution (Kuhn, 1962; Storer, 1966).

Stephens's prescription was to relax and enjoy the powerful and pervasive forces which apparently work well apart from any direct control. Dubin and Taveggia prescribed new research directions to find the commonalities among distinctive technologies of teaching; and to develop models of the teaching-learning situation. Although the two prescriptions may produce a satisfied and busy emperor, other evidence discounts the likelihood that they will provide him clothes.

Investigations in fields other than education suggest that research is at best an indirect route to improvement. The Project Hindsight study of utility of research in the military (Sherwin and Isenson, 1967) showed, for example, that it is unusual for random, disconnected fragments of scientific knowledge to be applied rapidly. The Project Hindsight investigators (Sherwin and Isenson, 1967, p. 1577) found that undirected science typically enters the world of practical affairs not by a route from the laboratory into life, but "in the compressed, highly organized form of well-established, clearly expressed general theory, or in the evaluated, ordered knowledge of handbooks, textbooks, and university courses" This compression and utilization takes twenty to fifty years. A subsequent analysis sponsored by the National Science Foundation (Thompson, 1969, p. 374) indicated that basic research is important but "about ninety percent of the basic research behind an innovation has been accomplished a decade before development of the innovation."

Fields which have looked to science as a source of improvement have devoted considerable effort to supplementing science with technology. "... To equate technology and science is wrong. Science is the quest for more or less abstract knowledge, whereas technology is the application of organized knowledge to help solve problems in our society" (Wiesner, 1966, p. 11). "... Technological knowledge is a set of techniques, each defined as a set of actions and decision rules guiding their sequential application that man has learned will generally lead to a predictable (and sometimes desirable) outcome under certain specified circumstances" (Nel-

son, Peck, and Kalachek, 1967, p. 8). The activity resulting from the application of such decision rules constitutes development—"... the systematic use of scientific knowledge directed toward the production of useful materials, devices, systems, or methods, including design and development of prototypes and processes" (National Science Foundation, 1965).

A considerable corpus of methodological literature on technology and development is available. The annotated bibliography prepared by Caldwell (1968) is an excellent traveler's guide to the field. Schon (1967) provided an easy overview to compare technological with scientific thinking and operations.

### *Comprehensive or Conjunctive?*

Most educational researchers have viewed methodological contributions as a means of creating a comprehensive, coherent body of knowledge about the phenomenon termed education. Storer (1968) presented a persuasive argument that such a goal is inadvisable, if not impossible. Rather than a discipline "organized around the investigation of empirical phenomena which seem to hang together regardless of men's individual or collective interests," educational research appears to be a conjunctive domain "defined so as to include all the phenomena which are involved in a specific social need or problem" (Storer, 1968, p. 5). If a curriculum researcher were to adopt this reasoning he would defer any attempt to codify a comprehensive curriculum research methodology and would look rather to the social sciences—psychology, sociology, anthropology, economics, and political science—and to the descriptive social disciplines such as geography, linguistics, and history. Operationally this would not represent a great deviation from prevailing practice; it would involve little more than extending the current perspective to fields other than psychology.

Fortunately, members of a number of conjoining disciplines have been developing and articulating methodology in a form that can be comprehended and used by persons with only journeyman curriculum-specialist training. The following examples are intended to suggest, not exhaust, the potential.

*Organizational research.* Curriculum research has given little direct attention to the organizational context in which instructional activities occur, other than to wail about the inadequacy of the age-grade structured bureaucracy. A large literature has accrued to help move the design of organizations from an art to a science (Likert, 1967; Thompson, 1966).

*Program budgeting.* Curriculum costs. The concept of program planning and budgeting systems which was introduced into defense planning in the McNamara era and which by Presidential directive is now being applied to all federal departments has considerable heuristic potential in

curriculum research. The system was well described by Novick (1965) and its historical perspective arrayed by Hitch (1965).

*Values and success indicants.* Curriculum researchers have often been hamstrung by intangibles which have been sensed but not defined. The Delphi technique (Helmer, 1966), which deals with values in terms of consensual validity, does not solve all of the philosophical questions which arise, but it does have the merit of rendering the ham and the strings visible. The technique is a means of concurrently reducing the ambiguity in the definition of a specified item and converging the personal judgments of specified individuals concerning the item. The analysis operates in real-time in contrast to conventional statistical approaches.

Bauer's (1966) discursive analysis of social indicators of "success" in the NASA context exemplifies what can be done with a large-scale and messy publicly-financed enterprise of much less direct immediate social value than education.

*Policy formulation and planning.* Techniques have been developed ranging from the macro to the micro level. At the broadest planning level, a number of the dilemmas of policy making were identified, if not resolved, in nine papers collected and edited by Bauer and Gergen (1968). At an intermediate level, specific forecasting techniques were described in a compendium edited by Bright (1968). At the micro-level, CPM-PERT sequenced event planning techniques were described by a number of authors (for example, Cook, 1966) and applied to large scale research planning by Carrese and Baker (1967).

*Politics of science and education.* Educational researchers in general have eschewed politics, regarding it as outside their professional domain. This provincial arrogance currently appears to be professionally unwarranted from the perspective of either politics or education. Three analytic areas of a political nature appear particularly relevant to curriculum research. First, the concept of organizational intelligence (Wilensky, 1967) utilizes the military rather than the psychological meaning of the term *intelligence*. Careful consideration of Wilensky's analysis should contribute to the curriculum researcher's intelligence in both the military and the psychological sense. Second, the politics of research as a social enterprise was treated in analytic surveys by Barber (1966) and by Greenberg (1967). The analyses indicated that some researchers are politicians all of the time, and all researchers are politicians some of the time; the relationship does not appear to be reversible for politicians. Third, the politics of education per se is beginning to draw the attention of political and sociological researchers (Innaccone, 1967). Political analyses of events in education reflect sources of variance which might well be included in the purview of the curriculum researcher. See, for instance, Bailey and Mosher (1968) for an analysis—on the federal level—of the sequences of events surrounding the Elementary and Secondary Education Act of 1965 which had impact



on the U.S. Office of Education, and Rogers's (1968) analysis—at the large urban district level—of the desegregation efforts of the New York City Schools.

### *Perseveration or Change?*

The response of the classical researcher to the foregoing may well be, "It has preciously little to do with what I call curriculum research." Precisely. While such a contention is presently irrefutable, this situation is not likely to hold for the future. Methodology accrues independently of the views or vested interests of self- or professionally-recognized methodologists, and its use likewise occurs independently of such views and interests.

### *Bibliography*

- Bailey, S. K. and Mosher, E. K. *ESEA: The Office of Education Administers a Law*. New York: Syracuse Univ. Press, 1968.
- Bakan, David. *On Method: Toward a Reconstruction of Psychological Investigation*. San Francisco: Jossey-Bass, 1967.
- Barber, Richard J. *The Politics of Research*. Washington, D. C.: Public Affairs Press, 1966.
- Bauer, Raymond A. (editor). *Social Indicators*. Cambridge, Mass.: M.I.T. Press, 1966.
- Bauer, Raymond A. and Gergen, Kenneth J. (editors). *The Study of Policy Formation*. New York: The Free Press, 1968.
- Berrien, Kenneth F. *General and Social Systems*. New Brunswick, N. J.: Rutgers Univ. Press, 1968.
- Bright, James R. *Technological Forecasting for Industry and Government: Methods of Application*. Englewood Cliffs, N. J.: Prentice Hall, 1968.
- Buckley, Walter (editor). *Modern Systems Research for the Behavioral Scientist*. Chicago: Ill.: Aldine, 1968.
- Caldwell, Lynton K. (editor). *Science, Technology, and Public Policy: A Bibliography I*. Bloomington: Department of Government, Indiana Univ., 1968. P. 492.
- Carrese, Louis M. and Baker, Carl G. The Convergence Technique: A Method for the Planning and Programming of Research Efforts. *Management Science* 13: B420-B438; 1967.
- Churchman, C. West. *The Systems Approach*. New York: Delacorte Press, 1968.
- Cook, Desmond L. *Program Evaluation and Review Techniques*. OE-12024. Cooperative Research Monograph No. 17. Washington, D.C.: Supt. of Documents, U. S. Government Printing Office, 1966.
- Coombs, Philip H. *The World Educational Crisis: A Systems Analysis*. New York: Oxford Univ. Press, 1968.
- Dubin, Robert and Taveggia, Thomas C. *The Teaching-Learning Paradox*. Eugene: Univ. of Ore., 1968.
- Educational Testing Service. *Proceedings of the 1967 Invitational Conference on Testing Problems*. Princeton, N. J.: The Service, 1968.
- Glass, Gene V. *Some Observations on Training Educational Researchers*. Research Paper No. 22. Boulder: Lab of Educ. Res., Univ. of Colo., Nov. 22, 1968.
- Goodlad, John I. Thought, Invention, and Research in the Advancement of Education. *The Educational Forum* 33: 7-18; 1968.
- Greenberg, Daniel S. *The Politics of Pure Science*. New York: The New American Library, 1967.
- Helmer, Olaf. *Social Technology*. New York: Basic Books, 1966.
- Hitch, Charles. *Decision-Making for Defense*. Berkeley: Univ. of Calif. Press, 1965.
- Innaccione, L. *Politics in Education*. New York: N. Y. Center for Applied Res. in Educ., 1967.
- Kuhn, Thomas S. *The Structure of Scientific Revolutions*. Chicago, Ill.: Univ. of Chicago Press, 1962.



- Likert, Rensis. *The Human Organization: Its Management and Value*. New York: McGraw-Hill, 1967.
- National Science Foundation. *Federal Funds for Research, Development, and Other Scientific Activities*. No. 19. Washington, D.C.: The Foundation, 1965. Pp. 56-7.
- Nelson, Richard R.; Peck, Merton J.; and Kalachek, Edward D. *Technology, Economic Growth and Public Policy*. Washington, D.C.: The Brookings Institution, 1967.
- Novick, David (editor). *Program Budgeting . . . Program Analysis and the Federal Budget*. Santa Monica, Calif.: The Rand Corp., 1965.
- Platt, John R. *Strong Inference*. *Science* 146: 347-52; 1964.
- Rogers, David. *110 Livingstone Street: Politics and Bureaucracy in the New York City School System*. New York: Random House, 1968.
- Schon, Donald A. *Technology and Change: The New Heraclitus*. New York: Delacorte Press, 1967.
- Schutz, Richard E. and Baker, Robert L. *The Experimental Analysis of Behavior. Psychology in the Schools* 5: 240-47; 1968.
- Sherwin, Chalmers W. and Isenson, Raymond S. *Project Hindsight*. *Science* 156: 1571-77; 1967.
- Sidman, Murray. *Tactics of Scientific Research*. New York: Basic Books, 1960.
- Simon, Herbert A. *The Architecture of Complexity. Proceedings of the American Philosophical Society*. Philadelphia, Pa.: the Society, 1962. Pp. 106, 467-482.
- Smith, Louis M. and Keith, Pat M. *Fantasy and Reality in Language of the Technology. Educational Technology* 8: 5-9; 1968.
- Stanley, Julian (editor). *Improving Experimental Design and Statistical Analysis*. Chicago, Ill.: Rand McNally, 1966.
- Stephens, John. *The Process of Schooling: A Psychological Examination*. New York: Holt, Rinehart and Winston, 1966.
- Storer, Norman W. *The Social System of Science*. New York: Holt, Rinehart and Winston, 1966.
- Storer, Norman W. *The Organization and Differentiation of the Scientific Community: Basic Disciplines, Applied Research, and Conjunctive Domains. Report of Colloquium on Improving the Social and Communication Mechanisms of Educational Research*. Washington, D.C.: Amer. Educ. Res. Assoc., 1969.
- Thompson, James D. (editor). *Approaches to Organizational Design*. Second edition, Pittsburgh, Pa.: Univ. of Pittsburgh Press, 1966.
- Thompson, Peter. *Traces: Basic Research Links to Technology Appraised*. *Science* 163: 374-75; 1969.
- Weisner, Jerome B. *Technology and Innovation. Technological Innovation and Society*. (Edited by Dean Morse and Aaron Warner.) New York: Columbia Univ. Press, 1966. Pp. 11-26.
- Wilensky, Harold L. *Organizational Intelligence: Knowledge and Policy in Government and Industry*. New York: Basic Books, 1967.
- Wittrock, Merlin and Wiley, David. *The Evaluation of Instruction: Issues and Problems*. New York: Holt, Rinehart and Winston, in press.

## AUTHOR

SCHUTZ, RICHARD E. Address: Southwest Regional Laboratory for Educational Research and Development Title: Director Age: 39 Degrees: B.A., Univ. of Calif., Los Angeles; M.A., Univ. of Calif., Los Angeles; Ph.D., Columbia Univ. Specialization: Research Management and Educational Psychology.

## 6: CURRICULUM: STATE OF THE FIELD

John I. Goodlad

*University of California, Los Angeles  
and Institute for Development of Educational  
Activities, Inc.*

In this chapter I seek to do three things: 1) analyze chapters 1-5 of this issue of the *Review* with a view to determining recent trends in curriculum research and development; 2) identify, partly from this analysis, continuing problems and issues in the curriculum field; and 3) appraise curriculum as a field of inquiry, giving some attention to future directions. In the concluding paragraphs of this chapter, I use the June 1960 issue of the *Review* as a touchstone in commenting on the past decade. The three purposes specified above are dealt with synthetically, not serially, and more or less in essay form.

Assuming that preceding chapters are reasonably comprehensive, one must conclude that increased concern for methodology characterizes current curriculum inquiry. It is significant that an entire chapter (Schutz) is devoted to methodological problems and issues. In two other chapters, Popham and Baker are much preoccupied with methodological problems in curriculum research, and McNeil concludes his chapter with recognition of a new emphasis on experimental work within the field. In this regard, it is gratifying to note that Baker, Popham and Schutz are highly critical of those comparative studies, which are all too common in educational circles, that purport to measure the effects of supposedly differentiated educational treatments when such treatments either are not described or are fictitious. The results of a clutch of such studies often are used, quite misleadingly, to commend or discourage particular educational practices.

Unfortunately, since this and most reviews of educational research tend to emphasize findings, it is impossible to know when to attach significance to what is reported. The addition of a methodological chapter in a review of curriculum developments is a welcome innovation. But curriculum inquiry—nay, educational inquiry—would profit immeasurably from truly critical reviews of both process and product in research. Such reviews might well include only a fraction of the publications cited here, but would examine them in depth.

Allow me to add one highly subjective comment regarding methodology. Curriculum researchers appear to be hung up on a limited repertoire of

methodologies borrowed primarily from psychology although, admittedly, preceding chapters include a few studies employing methods from other disciplines. The field of curriculum is still in its infancy and has not yet struck out boldly in an effort to employ, adapt, and invent methodologies suited to the peculiar character of the problems at hand. McNeil concludes his chapter with the puzzling query: "If researchers and practitioners cannot exercise control over a societal force and regulate it in the interest of better curriculum, then why bother to identify it as a force?" His preceding comment about the difficulty of manipulating a societal force is an understatement to say the least! But is not a prior effort necessary, an effort designed to get at the existential character of curriculum activity where it occurs? Objective field study—not requiring hypotheses, controls, and pretests (see Baker)—might well generate significant hypotheses for subsequent test in field experiments. These efforts would then lead to simulated studies with properties attributed to variables programmed for and manipulated by computers. Such has been the course of advancement in other fields.

Turning from the process to the content side of educational science, one is struck in preceding chapters with the paucity of ordered "findings" from curriculum research—findings in the sense either of scientific conclusions from cumulative inquiry or of tested guidelines for curriculum decisions. There does not seem to be a sizeable community of curriculum scholars who have staked out domains of inquiry with such clarity that successive studies are integrated into a larger whole, gaps identified, and new studies initiated. Nor does there appear to be a companion group of curriculum engineers aware of and drawing upon an established body of professional lore based on research. Schutz's succinct conclusion that "Education appears to be in a prescientific state rather than in a state of scientific revolution" probably applies as much to curriculum as to other subdivisions of education.

Probing this matter of findings, McNeil attests to the relative popularity of curriculum research pertaining to the *how* over that pertaining to the *what* of instruction. My reading of preceding chapters leads precisely to this conclusion. And it is a puzzling and disturbing one. Where, for example, are studies in the subject-fields? For more than a decade, millions of dollars have been spent each year in restructuring the subjects of elementary and secondary education. And yet this effort is scarcely mentioned, even in the historical chapter (Bellack). Perhaps this is because the so-called curriculum reform movement has been characterized by the refinement of curricula based on certain prior principles rather than by research. Perhaps such decision-oriented research as did accompany the movement was to be included in the missing chapter referred to by Popham in his introductory comments to this issue of the *Review*. Or, perhaps the omission simply represents an arbitrary cutting of the curriculum cake, with studies in the several subject-fields left to other issues of the *Review*.



Whatever the reasons were, this omission of studies in the subject-matter fields probably would produce in the non-curriculum specialist reading these chapters a rather strange feeling of unreality. The contents, for him, might be scarcely discernible as "curriculum." After all, is not *the* curriculum a specification of subject-matter to be learned? Therefore, should not curriculum research concern itself with the quality of such subject-matter, its relation to conceptions of knowledge and knowing, its impact on the value systems of learners, and like matters—all dealt with obliquely, if at all, in preceding chapters?

The authors of these chapters represent a relatively new and growing group of scholars—theorists and researchers in the field of curriculum. They are primarily (but not exclusively) concerned with advancing their chosen field of scholarship and only secondarily with curricula as they exist in thousands of schools and colleges. Consequently, it is neither surprising nor necessarily a bad thing that what they write about has no one-to-one relationship to what students and even teachers are above their ears in each day. If the abstract categories of research and discourse with which those scholars deal bear no identifiable relationship to the existential phenomena called curricula, then there is, indeed, cause for concern.

In addition to studies readily classifiable in the currently-useful categories of the scholars, there must be naturalistic studies of "the way it is." Goodlad (1968) and, more eloquently, Schwab (1969) appealed for such inquiry, particularly for the development of a curriculum theory with truly useful explanatory and predictive power. We simply do not know very much about the way it is at any level (societal, institutional, or instructional) of curriculum development or decision-making. Curriculum researchers have not identified the gates through which ideas pass nor what happens to them on the way to the classroom (Bellack and McNeil). Researchers know little about what happens in the classroom: how those carefully-developed materials are used or if they are used at all; how conflicts between the ideological curriculum of materials and the mental curriculum of teachers are reconciled; what reaches and attracts the student and what does not; and on and on. There is no doubt about the need for better curriculum experiments, as several chapter authors point out. But paralleling these must be equally well but differently designed studies seeking to explain the nature of what reasonably can be called "the curriculum," what affects its formulation, and what it affects and how. The mode in educational research has been the reduction (via multivariate analysis) of the number of variables, when, in the words of Easley (1967), "... perhaps what is most needed is to go out and find more relevant ones."

At the classroom or instructional level of curriculum, the work of Bellack and Smith is well covered (see McNeil). But the even more naturalistic studies of Jackson (1968) and of Smith and Geoffrey (1968) are not discussed at all. Perhaps this is because the latter pair of studies



encompasses much more than is encompassed by any reasonable definition of curriculum; perhaps curriculum as a field of study has not yet adequately legitimized such studies.

Another area of omission should be noted. Reviews of curriculum work in the urban ghetto and with minority group children are few in number. Admittedly, although much of this work has been action-oriented, there does exist a growing body of research, much of it to be found in doctoral dissertations and government reports. The bulk of the research is at the early childhood level, a level perhaps not yet institutionalized in curriculum inquiry. So far as this issue of the *Review* reveals, very little is known about the differences in degree and kind of problems encountered by curriculum researchers and developers extending their interests into urban environments.

To this point, I have separated for discussion purposes the two domains of method and subject-matter in curriculum research, both essential to any meaningful discourse on educational science or practice. In a recent publication of the National Academy of Education (1969), the authors trace what they term "chains of inquiry" in the development of educational thought and practice. Are any such chains apparent in the field of curriculum? Perhaps the word "chains" connotes unjustified strengths as far as curriculum inquiry is concerned, but two closely-linked threads are identifiable: the formulation and use of educational objectives and the systematic arrangement of instructional stimuli. Both have a substantial lineage, as Bellack points out. It is interesting to note that preoccupation with educational objectives has gone through several stages of conceptual and taxonomical analysis (see Bellack) into a stage of experimentation (see Popham). This progression parallels progress in most sciences. Arranging instructional stimuli has been linked for decades with the textbook industry; now this curriculum problem and the problem of formulating educational objectives are closely linked with both hardware and software in the publishing field. Ties with industry are familiar linkages in the development of any science. Whether these ties delay or enhance the advancement of a science, however, is not at all clear.

We must not fall into the error of thinking that educational science—unlike other sciences—will advance in an orderly, cumulative fashion. A new theory "... requires the reconstruction of prior thinking and the reevaluation of prior fact, an intrinsically revolutionary process that is seldom completed by a single man and never overnight," (Kuhn, 1962, p. 7). That a reaction has begun against the specification of precise, behavioral objectives and programed learning sets attests to some maturity in these aspects of curriculum. The technical and, to a lesser degree, philosophical disagreements underlying this curriculum dispute are documented by both Bellack and McNeil, and most of the authors in this issue give at least lip-service to what appears to be a rather polite intellectual disagreement. One wonders, however, whether the deeper conflicts are

generally recognized. Ends/means manipulations in curriculum are the handmaidens of behaviorism. Behaviorism and rediscovered and reinterpreted humanism are rubbing against each other more abrasively. Is an embryonic revolution in the hitherto rather sleepy field of curriculum about to be born of the social revolution of our time? Similar occurrences are not without precedent in the history of science.

The preceding reviews suggest that curriculum evaluation, too, may be moving out of a pre-scientific state. This aspect of curriculum lay virtually dormant for two decades following a flurry of activity in the late 1930's and early 1940's (McNeil). There is a flurry of activity once more—enough, for example, to generate a recent yearbook of the National Society for the Study of Education (1969). But, as Baker points out, there is little focus. Unfortunately, the unsteadiness of the field simultaneously fosters easy expertise and thwarts critical attack.

The contribution of evaluation to curriculum development has been restricted by a model which has extended very little beyond conceptions of testing. Tests have been designed to measure the attainment of certain curricular goals and the results obtained from them have then been used to make evaluative statements regarding students' learning or teachers' teaching. This simplistic model assumes that the efforts of teachers and students are the prime causal factors in the attainment of educational objectives. Researchers know, however, that there are other causal factors; evaluation should help to identify discrepancies or inadequacies in any of several factors included in a comprehensive model. Some recent evaluative efforts reflect cognizance of this need. Among them is the Discrepancy Evaluation Model being used in the city of Pittsburgh (Provus, 1969). Evaluation must be designed to determine the relationship not only between stated goals and the degree of their attainment but also between an initial conception of the practice designed to achieve them and the practice actually developed. Such diagnostic information is definitely more helpful than knowledge about attainment of objectives alone.

One development of international proportions promises to move this kind of evaluative thinking along. More than a dozen nations are collaborating on cross-national studies of educational achievement (Husén, 1967). Studies in mathematics have been reported and studies in other fields are underway. Japan, Sweden, and the United States, all considered to be educationally advanced, differed markedly in their standings in mathematics achievement; Japan headed the list and Sweden and the United States ranked near the bottom. Why did the U.S. and Sweden do poorly? The simple answer is that mathematics is poorly taught or inadequately emphasized in the schools of these two countries. Although this could be the correct answer, a more comprehensive model of explanatory factors would turn up more possibilities and protect against myopic judgments. The answer could also lie in the educational aspirations of Japanese

boys and girls, in parental reinforcement of school activities, or in other aspects of Japanese culture.

Unfortunately, the major components of a causal model of learning, let alone the respective contribution of each component, are unknown. How much of the whole do we assign to family context, to teachers, to the curriculum? The answers will be a long time in coming. The preceding reviews of research in this issue suggest that the curriculum field is not very advanced in the conceptualizing and theory building needed to find the answers. The trees are receiving increasingly systematic attention but it is difficult to know whether analysis of them will add up to a better understanding of the forest.

At the outset, I indicated that the June 1960 issue of the *Review* would be used as a beginning point for appraising curriculum research and development over a period of nearly ten years. Chapter One of that issue, "Curriculum: State of the Field," provides a good basis for assessing the progress that has been made. Chapter Three of the 1960 issue identified an emphasis on redefining and restructuring content, citing a series of large scale projects in mathematics, languages, and the physical and biological sciences. As I maintained earlier, there are only a few references to such developments in the present issue. Several possible explanations, including that of a missing chapter in this 1969 issue, are set forth on preceding pages. Nonetheless, such work continues on a large scale and has been extended to include the social studies, English, the arts, and health education. It was suggested in the 1960 issue that this heightened interest in content might counterbalance a forty-year trend toward the behavioristic in education. Assuming that the 1969 authors have been reasonably comprehensive, one must conclude, however, that the forty-year-old behavioristic cycle has reached the half-century mark.

Authors of the 1960 issue reported no studies establishing an actual relationship between increased clarification of educational objectives and improved discrimination in the selection of classroom learning opportunities for students and took special note of this fact. The current issue reports substantial evidence regarding this relationship (Popham). The earlier issue contained a plea for extending taxonomical analyses of educational objectives into psychomotor and affective realms of behavior. Taxonomies in both have since been developed (see McNeil). Still in the realm of educational objectives, the 1960 issue referred to the danger of seeing only one type of behavior in a statement of ends when in reality two or more may be involved. Current research offers little about the processes of derivation involved in the structural decomposition of an educational objective, but it presents a great deal about the refinement of a single type into precise, behavioral sub-objectives (Popham). In summary, in the past ten years little has been added to knowledge of deriving educational objectives; substantial progress has been made in stating and using educational objectives with precision.



The authors of the 1960 issue maintained that "'global' approaches to the establishment of relationships between curriculum planning processes and improved instruction and learning must be replaced by research studies more precisely isolating and comparing process-product factors." The research summarized in at least two current chapters (Popham and Baker) attests to progress here. However, two major problems are likely to be with us for some time. The first of these problems is methodological; it is well illustrated by both Baker and Schutz in their comments about clearly delineating a process before attempting to assess or attach significance to the product. The second problem is theoretical or conceptual and pertains to the difficulty of knowing the place of a small, manageable process-product equation within some larger frame of explanation.

One area of research which apparently was almost dormant in 1960 seems to be very much alive in 1969. This is the quest for how best to arrange material in a field for effective learning. Studies are cited in several of the chapters of this issue; many more might have been included, especially those pertaining to programing computer sequences. Research of this kind—in spelling, reading, and mathematics, in particular—was relatively popular in the 1920's and 1930's but faded noticeably in the 1940's and 1950's. It is back in mode again, at least with a small group of able researchers.

To quote again from the 1960 issue: "Theoretical constructs are needed from which research studies may be derived to demonstrate how values and expectations of individuals and groups find their ways through various channels of communication and political (conceived in the broadest sense) structures to influence curriculums." But, in 1969, both Bellack in his historical analysis and McNeil in his contemporary analysis point to the difficulties and limitations in ascribing specific influence on the curriculum either to persons or to forces not clearly identified with individuals. Awareness of the need to trace various factors thought to influence the curriculum as a step toward building predictive theory is more poignantly expressed in 1969 than it was in 1960. Growing awareness of a human need usually leads to corrective action and perhaps we can be more optimistic in 1969 than we were in 1960.

As far as theory *qua* theory in any realm of curriculum inquiry is concerned, however, the cupboard is almost bare. A bone to chew on is Beauchamp's *Curriculum Theory* (1968), a revised edition of a 1961 publication. However, Beauchamp did not even attempt to develop a curriculum theory. Instead, because of the lack of materials from which to build a theory at that time (which Beauchamp pointedly recognized), he was forced to content himself largely with a discourse on theory building.

In the conceptual realm, the 1960 issue included the following statement: "Conceptual systems which identify the major questions to be



answered in developing a curriculum must be rigorously formulated. The elements that tie these questions together in a system must be classified; subordinate questions must be identified and classified properly in relation to the major questions; sources of data to be used must be revealed in answering the questions posed by the system; and the relevance of data extracted from these sources must be suggested." As far as the major questions to be answered in developing a curriculum are concerned, most of the authors in 1960 and 1969 assume those set forth in 1950 by Ralph Tyler. In both issues a substantial portion of the research reviewed is organized around Tyler's questions. No other scheme has served in a similar way during this twenty-year period. Two authors, Baker and McNeil, refer to a more elaborate conceptual scheme (Goodlad, 1966), which includes the Tyler rationale, but there is as yet no evidence to suggest that it is serving to stimulate research or to organize thinking in the curriculum field. General theory and conceptualization in curriculum appear to have advanced very little during the last decade.

In brief summary, during the past decade significant progress has been made in the precise definition of curricular objectives, in the analysis of ends/means relationships, and in the effective ordering of stimuli for learning. Substantial progress has been made in extending both the understanding of the evaluative process and the use of evaluative data in diagnosing the possible causes of discrepancies between curricular expectations and curricular accomplishments. In the realm of explaining curricular realities, however, we appear to know little more in 1969 than we knew in 1960. Curricular theory with exploratory and predictive power is virtually non-existent. The most eloquent plea (Schwab, 1969) for correcting this condition directs our attention to the existential character of the curriculum: what it is, how it gets to be the way it is, and how it affects the people who partake of it.

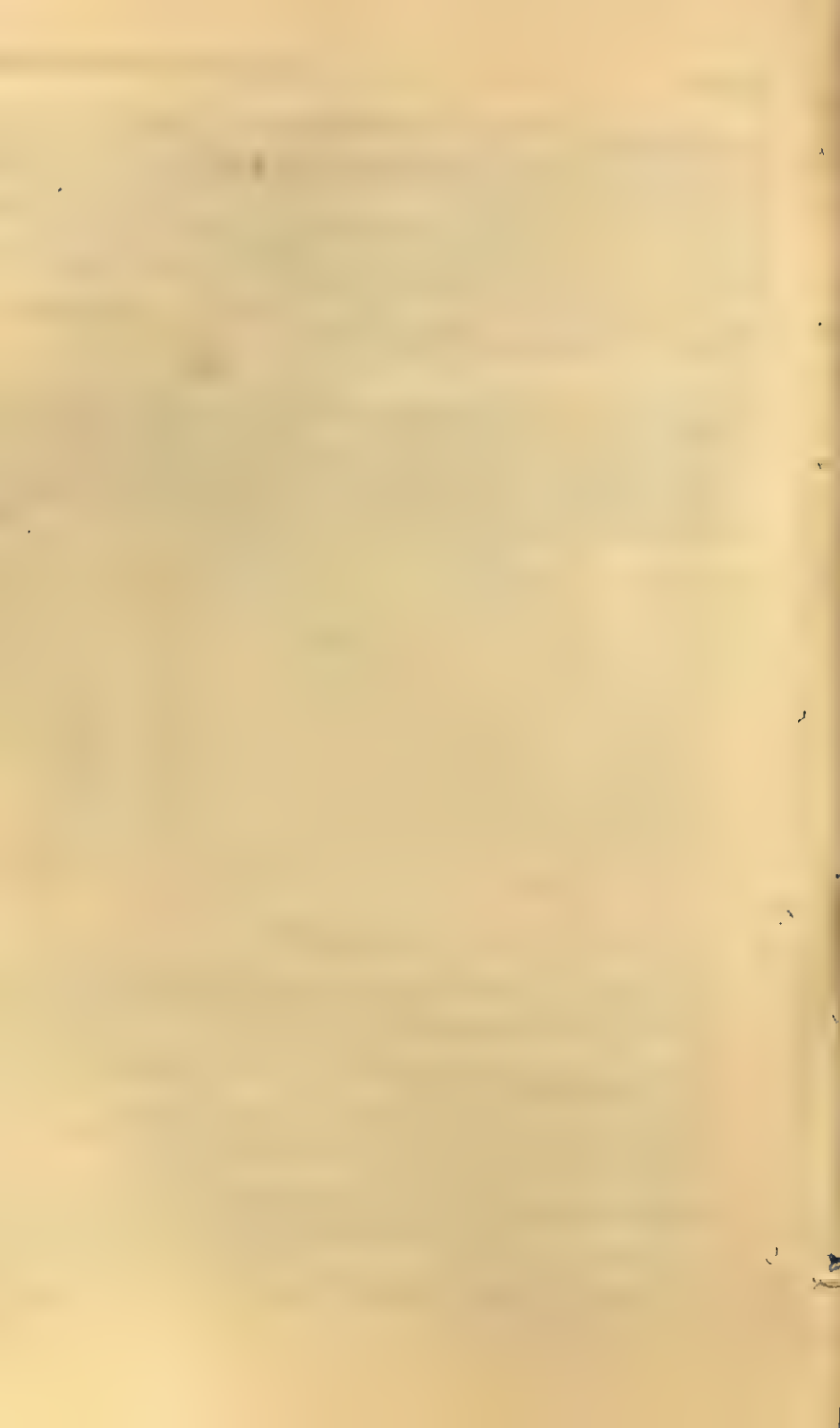
### Bibliography

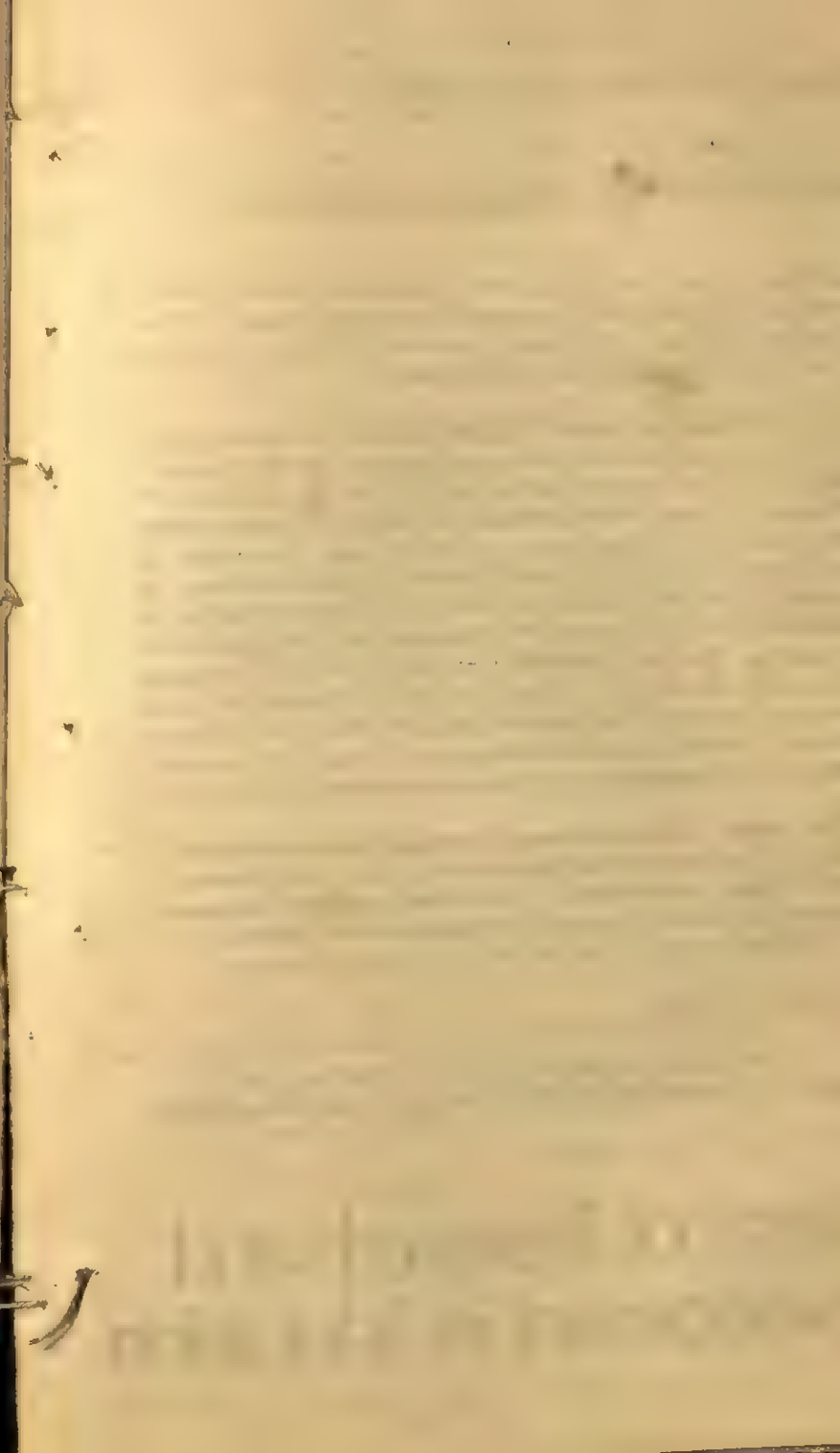
- American Educational Research Assoc. *Guidance and Counseling. Review of Educational Research* 30, No. 2: 1960.
- Beauchamp, George A. *Curriculum Theory* (second edition). Wilmette, Ill.: Kagg Press, 1968.
- Easley, John A., Jr. *Curriculum Development, Teacher Education, and Educational Research*. Unpublished paper. Urbana: Univ. of Ill., Curriculum Laboratory, 1967. (Mimeo.)
- Goodlad, John I. and Richter, Maurice N., Jr. *The Development of a Conceptual System for Dealing with Problems of Curriculum and Instruction*. Cooperative Research Project No. 454. Los Angeles: Univ. of Calif. and Institute for Development of Educational Activities, 1966.
- Goodlad, John I. *Thought, Invention, and Research in the Advancement of Education. The Educational Forum* 33, No. 1: 7-18; 1968.
- Husén, Thorsten (editor). *International Study of Achievement in Mathematics*, Vols. I and II. New York: John Wiley, 1967.
- Jackson, Philip W. *Life in Classrooms*. New York: Holt, Rinehart and Winston, 1968.
- Kuhn, Thomas S. *The Structure of Scientific Revolutions*. Chicago: Univ. of Chicago Press, 1962.

- National Academy of Education. *Disciplined Inquiry for Education*. New York: Macmillan, 1969.
- National Society for the Study of Education. *Educational Evaluation: New Roles, New Means*. (Edited by Ralph W. Tyler.) Sixty-eighth Yearbook, Part II. Chicago: Univ. of Chicago Press, 1969.
- Provus, Malcolm. *The Discrepancy Evaluation Model: An Approach to Local Program Improvement and Development*. Pittsburgh: Pittsburgh Public Schools, 1969.
- Schwab, Joseph J. *The Practical: A Language for Curriculum*. Unpublished Address, Annual Conference of the Amer. Educ. Res. Assoc., February, 1969. Chicago: Univ. of Chi., School of Education.
- Smith, Louis M. and Geoffrey, William. *The Complexities of an Urban Classroom*. New York: Holt, Rinehart and Winston, 1968.

## AUTHOR

GOODLAD, JOHN I. Addresses: University of California, Los Angeles and Institute for Development of Educational Activities, Inc. Titles: Dean, Graduate School of Education, UCLA, and Director, Research and Development Division, IDEA. Age: 48. Degrees: B.A., Univ. of British Columbia; M.A., Univ. of British Columbia; Ph.D., Univ. of Chicago; L.H.D., National College of Educ.; L.H.D., Univ. of Louisville. Specialization: Curriculum Theory and Practice, Study of Educational Change, Inquiries into Elementary School Patterns of Organization.







Volume 2 • Number 4 • Summer, 1969

PUBLIC INTEREST *vis a vis* EDUCATIONAL R & D

Proceedings of the June 16-17, 1969 Conference at Airlie House, Warrenton, Virginia jointly sponsored by *Saturday Review* and the *Journal of Research and Development in Education* with the cooperation of the U S Office of Education.

Table of Contents includes: "Political Realities of Educational R & D" by Norman J. Boyan, Associate Commissioner for Research, USOE; "The Institutional Base for R & D: What More is Needed?" by Francis S. Chase, Visiting Professor, College of Education, Ohio State University; "Significance of the Federal Investment in Educational R & D" by Stephen K. Bailey, Maxwell Professor of Political Science, Syracuse University; "The R & D Problem in the Behavioral and Social Sciences" by Ernest R. Hilgard, Professor of Psychology, Stanford University; reports from the long-term policy research centers by Thomas F. Green, Director, Syracuse University Research Corporation and Dr. Willis W. Harman, Director, Stanford Research Institute; and remarks by Assistant Secretary and Commissioner of Education James E. Allen.

The purpose of the Conference-convened for the benefit of editorial policy makers of national news media, and observers representing government, quasi-official and professional organizations-was to examine directions of educational R & D and the kinds of support required to improve education.

Single copy - \$2.00

Send orders to Business Office, 122 Baldwin Hall,  
College of Education, University of Georgia, Athens, Georgia 30601

# Journal of Research and Development in Education

Spring, 1969  
Elementary  
Teacher  
Education  
Models

Please enter my subscription to the  
Journal of Research and Development in Education Starting  
with ☐ Spring ☐ Summer ☐ Fall ☐ Winter  
for ONE year at \$7.00 per year

Please send me \_\_\_\_\_ copies of the ELEMENTARY TEACHER EDUCATION  
MODELS issue at \$2.00 each.

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip Code \_\_\_\_\_

Please send and bill me for the following back issues:

- ☐ Reading Research - Volume 2, Number 2
- ☐ Beginning Teachers and Certification - Volume 2, Number 1
- ☐ USOE-Funded R & D Centers: An Assessment - Volume 1, Number 4
- ☐ Early Childhood Education - Volume 1, Number 3
- ☐ Group Counseling - Volume 1, Number 2
- ☐ Mathematics Research - Volume 1, Number 1

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip Code \_\_\_\_\_

Please accept my order and bill me for the following future issues:

- ☐ The Public Interest *Vis-À-Vis* Educational R & D  
- Volume 2, Number 4
- ☐ Classroom Communications: Improving Children's Thinking  
- Volume 3, Number 1
- ☐ Regional Educational Laboratories: Past, Present and Future Perspectives  
- Volume 3, Number 2

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip Code \_\_\_\_\_

Place  
Stamp  
Here

Business Office  
Journal of Research and Development in Education  
122 Baldwin Hall  
College of Education  
University of Georgia  
Athens, Georgia 30601

Place  
Stamp  
Here

Business Office  
Journal of Research and Development in Education  
122 Baldwin Hall  
College of Education  
University of Georgia  
Athens, Georgia 30601

Place  
Stamp  
Here

Business Office  
Journal of Research and Development in Education  
122 Baldwin Hall  
College of Education  
University of Georgia

## FOREWORD

This issue of the Review has been organized to highlight topics of current significance in science and mathematics education as we perceive them. The organization departs from previous issues dealing with these fields; we hope that researchers will be tolerant of the change and that accommodation to the new format will require minimum adjustment.

The intention in this issue has been to focus on general trends and issues. Therefore, the reader will find few comprehensive and detailed reviews of specific investigations. While considerable research is cited, description is held to a minimum to sharpen the exposition and to highlight the biases of the individual authors.

Gene V Glass, the General Editor, has offered insightful guidance at every step in the preparation of this issue. His strong hand has kept the volume of material to manageable proportions. It has also improved the quality. We thank him. The assistance to the General Editor of Drs. Arthur L. White, Ohio State University, and Ronald D. Anderson, University of Colorado, is also gratefully acknowledged.

J. Myron Atkin  
Thomas A. Romberg  
*Issue Editors*





# 1: LEARNING STUDIES IN SCIENCE EDUCATION

Maurice Belanger

Harvard University

The usual textbook definition of learning is "change in behavior." If this definition is adopted as the search criterion for identifying pertinent items for review in the science education literature, several difficulties are immediately encountered. The conceptualization of learning in this way could be used to subsume studies in evaluation, for example in curriculum evaluation where presumably the curriculum serves as a means for some modification of behavior. Or studies in attitude change or any study where achievement is measured might be appropriately included. However, if an attempt is made to narrow the range by adopting a notion such as Ausubel's "Meaningful Verbal Learning" or Piaget's notion of learning, in the broad sense, as simply another term for intelligence, then very few studies in science education could be identified as within either domain. Thus, the immediate problem is deciding what to review.

Consulting the standard indices, such as *Education Index*, *Psychological Abstracts*, and *Dissertation Abstracts*, shows that there is a relatively standard set of categories ("concept learning," "problem solving," "discovery learning," "critical thinking," etc.) for classifying learning studies. Each category might be thought to represent a closely bound set of ideas, but further problems emerge as one begins to read the studies. Take as an example the category "discovery learning." In a hypothetical study conducted by Mr. X and clearly labeled a discovery study in the title, second and third grade pupils were taught a series of "discovery lessons on flowers." The pupils used programed directions written in the Initial Teaching Alphabet. The "Mr. X Test of Discovery Learning" was given as a pre- and posttest. The reader knows little about the test other than that it consists of twenty items constructed by Mr. X. After analysis of pre-post data, Mr. X reports that discovery learning was far superior to conventionally taught lessons. In another hypothetical study, by Mr. Y, college freshmen in an introductory geology class were engaged in a series of "discovery lessons" employing geological maps on field trips. An achievement test constructed by Mr. Y was given to classes using the discovery lessons and to a control class using only the text and lectures. Analysis of data revealed significant differences in achievement in favor of the experimental discovery class.

Studies of this type could be reported on at length, but the general reader probably would like to know whether or not the research on discovery learning in science leads to any trustworthy generalizations. Even if impeccable research designs are found in the studies—and since this is rarely the case it becomes all too easy simply to describe and then tear apart the studies on technical grounds alone—can it be asserted on the basis of the X and Y studies, or thirty other variations, that “discovery learning” is superior to some other variety of learning? What is the status of the construct “discovery learning”?

A host of issues concerning discovery learning (as an internal process or as a method of teaching) are presented in Shulman and Keislar's *Learning by Discovery* (1966). In that volume, David Hawkins raised severe doubts about the validity of “discovery learning” as a pedagogical construct. But even for researchers, Wittrock in the same volume stated, “the aging but still illusive learning by discovery hypothesis has outlived its usefulness to researchers” (p. 74).

Consider yet another category—“concept learning.” Wallace, in his book *Concept Growth and Education of the Child* (1965) reviewing British, French and American literature on concept learning, concluded after 198 pages: “It is a sad commentary on the effectiveness of our methods of enquiry that after some eighty years of psychological investigation and a discontinuous history of forty years of laboratory experiments, our fund of accepted knowledge on the subject of conceptualization comprises so little of consequence that it is hardly worth compiling.”

Thus, the dilemma facing a reviewer of learning studies goes beyond identification, critique and generalization and extends to the nature of the learning categories themselves. A modest probe into the meaning of these categories immediately leads to difficult philosophical questions about the nature of knowledge and knowing.

From a pool of 150 items on science learning, published between Fall 1964 and Winter 1969, I have selected items which are classified within the traditional categories of science learning research.

### *Concept Learning*

Papers on concept learning in science outnumber papers in other categories, probably because the learning of concepts is frequently stated as a major goal of science teaching. The following list provides a minimal guide to selected papers. I have identified author(s), grade level of the group of subjects, the concept domain, the methods used, measures, and a brief statement of conclusions. General comments are given after the list.

Adler (1966)

Grades: College undergraduates. Concept domain: Space, matter and energy. Method:

Students taught a general education course in science. No control group. *Measures:* Science Concept Test by author, SCAT, Ohio State "Psychological Test," number of years of high school science. *Conclusions:* Development of concepts of space, matter and energy related to previous student background in physical science and to verbal and mathematical aptitude. Men scored slightly better than women, but had also taken more physical science.

Anderson (1965)

*Grades:* 3 to 6. *Concept domain:* Formulation of mental models as a network of concepts; temperature and change of state of water; surface tension; mixture of alcohol and water. *Method:* Demonstration and interviews; presentation of mechanical mode. *Measures:* California Test of Mental Maturity; analysis of children's responses to demonstrations and questions; classifications of models by type. *Conclusions:* Fewer girls used atomistic models; consistency of explanation and ability to formulate mental models increased with age; frequency of atomistic models increased with age and also related to IQ.

Brusini (1966)

*Grades:* 3 to 8. *Concept domain:* Cycle of stream erosion. *Method:* Arrangement of picture sequence. *Measure:* Correct arrangement of picture sequence. *Conclusion:* Greater success in arranging picture sequence for older children.

Dallas (1968)

*Grades:* College juniors and seniors in teacher preparation. *Concept domain:* Concepts selected from NSTA 1964 curriculum committee. *Method:* Ten science concepts taught in two structured orders. No instruction to control group. *Measures:* "Test of Concept Application" constructed by author. *Conclusions:* No difference in gain scores between group taught by hierarchical sequence of concepts and group taught by nonhierarchical order; for experimental group greater gains than control group; gains not related to students mathematical or verbal aptitude, science background, college grades or level of dogmatism.

Dennis (1966)

*Grades:* 2 to 6, subdivided into three subgroups by IQ scores. *Concept domain:* Kinetic molecular theory. *Method:* Ten one-half hour lessons; films, demonstrations and experiments; no control group. *Measures:* Pre-post tests by authors. *Conclusions:* Significant gains for students in "normal" third grade and above. Gains related to IQ scores.

Harris and Lee (1966)

*Grades:* 4 to 6, subdivided into groups by IQ. *Concept domain:* Molecular-kinetic theory. *Method:* Taped instructions to experimental group. 15 lessons; Control group received no instruction. *Measures:* Written and oral pre-post tests designed by authors. *Conclusions:* Greater test gains for experimental group; gains related to IQ scores; 5th and 6th graders outperformed 4th graders.

Kantz (1968)

*Grades:* 1, 3, and 5. One group examined in 1950, another group in 1964. *Concept domain:* 68 terms and principles of physical science. *Method:* Demonstration and individual interviews. *Conclusions:* In general, the older child more correct in concepts of terms and principles; Children interviewed in 1964 more scientifically correct than children interviewed in 1950.

Kolb (1967)

*Grade:* 5. *Concept domain:* Static friction and a sequence in mathematics (ratio and line graphs.) *Method:* Two science exercises designed on basis of AAAS Elementary Science Program; one group received special instruction in mathematical sequence designed by author, other group studied traditional mathematics text in use. *Measures:* Measure I, achievement in mathematics; Measure II, to assess behaviors as stated in science lessons. Measures I and II used as pre-post tests. *Conclusions:* Groups studying the special math sequence scored higher on science achievement measures than group studying regular mathematics text.

Pella and Ziegler (1967)

*Grades:* 4 to 6. *Concept domain:* Four concepts of kinetic molecular theory. *Method:* Comparison of children classified as "modelers" and "non-modelers" on basis of pretest; Two treatments (1) instruction with static models, (2) instruction with



dynamic models; Control group received no instruction. *Measures*: Pretest to measure use of mental models; Lorge-Thorndike IQ Test; Stanford Achievement (Science Scores). *Conclusions*: Use of acceptable models increased with grade level; Group of "modelers" outscored "non-modelers" within the same instructional method; Gains not related to IQ or prior achievement; No differences in gains in the use of static models vs. dynamic models in instruction.

Thier (1965)

*Grade*: 1. *Concept domain*: Concepts of matter, 6 categories used. *Method*: Demonstration and interview by standard protocol. Experimental group had studied SCIS unit "Material Objects"; control group had not studied "Material Objects." *Measures*: Analysis of children's responses to demonstrations and interview protocol questions. *Conclusions*: Children in experimental group more able to describe objects by their properties, more capable in descriptions of similarities and differences, and also in observations of an experiment and in the use of observations to describe accurately what had happened.

Trent (1965)

*Grades*: High school. *Concept domain*: "Understanding Science." *Method*: Experimental group taught a traditional physics course. *Measures*: "Test On Understanding Science" by Klopfer and Cooley; Otis Quick Scoring Mental Ability Test. *Conclusion*: When students' prior understanding of science and mental ability was controlled, no significant difference between gains for PSSC students and those in traditional physics course.

In the 11 papers listed, grade level ranged from first grade to college seniors; there was almost no overlap in the concept areas investigated. These two factors alone make generalizations risky. With the exception of Thier's study, none could be replicated because the tests designed by the investigators were not provided and the methods used in experimental treatments were described in such general terms that it is difficult to determine what was done. Perhaps the only major reliable outcome of these studies is that science concept learning is directly related to age and IQ. Only such gross statements are possible because most of the studies made only gross comparisons. Detailed probing, fine analysis, and even item analyses of tests are absent. It is very likely that future studies on science concept learning which employ single scores on tests as the major mode of analysis will continue to provide only the crudest information. Much more promising than simplistic comparative experimental designs is the use of interviews that provide sufficiently rich information to enable the investigator to generate tentative hypotheses beyond the commonsense level of generalization.

The studies by Thier, Pella and Ziegler, and Anderson are extremely promising beginnings in the use of interview techniques for the study of science concepts. The attempt to formulate patterns of concepts (models) seems to be an effective strategy for structuring the mass of information that interviews provided. These authors stated they employed Piaget-type interviews in the sense that a child was shown concrete objects or phenomena and then asked to respond to such aspects as description, prediction, or explanation. Although this procedure does contain parallels to Piaget interviews, there is an appreciable difference. In the interviews by the Genevans, the procedures are specifically focused on the probing of hypotheses contained within a rich theoretical structure.

In the nearly total absence of a theoretical framework for the study of science concepts, studies attempting to pit ad hoc hypotheses against each other will probably continue only to demonstrate the obvious or to add to the already confusing list of findings from "method X vs. method Y" investigations.

*Additional References:* Novak (1965); Pella and Carey (1967); Pella and Strauss (1967); Pella and Triesenberg (1967); Pella and Voelker (1967); Schaaf (1965).

### *Problem Solving*

The category of problem solving is quite ambiguous. Sometimes *problem solving* refers to a generalized method for seeking answers to questions within a discipline, but at other times it refers to an interdisciplinary decision-making process. A theoretical analysis of these two referents was the subject of a dissertation by Grant (1965). Cunningham's (1966a) excellent review dealt with rigidity in children's problem solving. More general papers on problem solving are those by Anderson and Weigand (1967), Anderson (1967), Mahan (1967), and Aylesworth (1965). The Mahan and Aylesworth papers reverted to the consideration of problem solving as a variant of scientific methods.

O'Toole (1966), Inventasch (1968) and Anderson (1966) compared modes of problem solving. In none of their studies was a significant difference on the dependent variables found in the comparisons of the two methods. The comparison of one general method of problem solving with another is a common strategy in research studies. In contrast, the work of Butts and his colleagues probed in considerable detail the processes that individuals use when challenged with a variety of science problems. This in-depth probing in the work of Butts is in stark contrast to the other studies listed below:

#### *Anderson (1967)*

*Grades:* College students. *Method:* Problem-solving instruction; 7 brainstorming sessions; control group received no instruction. *Measures:* Instrument by author; Nelson-Denny Reading Test; Watson-Glaser Critical Thinking Appraisal. *Conclusions:* No differences between problem-solving training group and control; significant correlations of test scores on Problem-Solving Test and Reading Scores, and problem solving with Watson-Glaser Test.

#### *Butts (1965)*

*Grade:* College seniors. *Method:* A problem-solving test was designed in which a problem was presented and the subject was enabled to select the kinds and amounts of information he believed would enable him to solve the problem; chains of reasoning by subject can then be examined. *Measures:* STEP (Science); problem-solving test by investigator; TAB: format for analysis of problem-solving behavior. (See Butts and Jones, 1967.) *Conclusion:* Very low correlations between Problem-Solving Test and STEP test.

#### *Butts and Jones (1966)*

*Grade:* 6. *Method:* Problem presentation using materials from Suchman's Inquiry Training; child asked to predict, explain, then given questions that might aid in solving the problem; child selected questions and was provided with information

and again asked to explain; transfer task then given to child. *Measures*: STEP (Science); TAB: Inventory of Science Progress by investigator. *Conclusions*: Positive relationship between inquiry training and problem-solving behavior; no concept transfer, or differences in recall of factual information; Not related to IQ, age, factual knowledge, or sex.

Inventasch (1968)

*Grade*: 7. *Method*: Teacher-directed vs. self-directed problem-solving techniques. *Measures*: STEP tests; Test on Understanding Science; Cooperative Science Tests. *Conclusions*: No difference in two methods; no differences by previous academic achievement.

O'Toole (1966)

*Grade*: 5. *Method*: Use of modified AAAS Science and science textbook; individualized instruction vs. teacher-centered methods in problem solving. *Measures*: Stanford Achievement Test; STEP (Science); SRA "What I Like to Do" and test by investigator; Self-concept check test. *Conclusions*: No difference by two methods for STEP test, achievement test by investigator, self-concept, interest, problem-solving abilities; gains in ability to identify hypotheses and problems.

Roberts (1964)

*Grade*: 9. *Method*: Students scoring in the upper third of the Science STEP test defined as "problem solvers," while students in lower third defined as "non-problem solvers"; students given ten problems from an alternate form of STEP and asked to "think aloud" while solving problems. *Measures*: STEP (Science) and analysis of student protocols by Problem-Solving Checklist. *Conclusions*: "Problem solvers" more confident, possessed more knowledge, more able in defining problem, picking out hypotheses, finding relevant material, using ideas, operated at "higher level of thought"; No difference in time taken to solve the problems.

As with concept learning studies, comparing one method of problem solving with another seems to produce few consistent differences as indicated by the measures used. It is again very difficult to generalize due to the variety of ages of subjects, problems presented and measures used. The work of Butts seems to have been the most promising in this area during the period under review.

### *Inductive-Deductive Learning*

Studies in this category are labeled with a variety of terms—*inductive-deductive* when comparing two methods and sometimes as *discovery-didactic*. Studies by Scheffler (1965), Tanner (1968, and Coulter (1966) employed different methods in which one mode of instruction was compared with another. No differences were found in different treatment groups in any of these studies. The study by Menefee (1965) focused on inference-making by elementary school children rather than on a comparison of methods. Guided discovery could also be placed in this general category. Studies by Salstrom (1967) and Thomas (1968) indicated better test performance for groups in guided discovery situations as compared to groups in other types of learning situations. The studies by Rainey (1965) and Lansdown and Dietz (1965) indicated superior performance by subjects in free discovery situations; these studies also provided a number of subjective reports based on the investigators' observations of classroom performances. Such commentary indicated that performance on tests may not indicate differences in free vs. guided discovery measured outcomes,



but there may be a number of other student behaviors that are enhanced by free discovery procedures.

### INDUCTIVE-DEDUCTIVE LEARNING

Coulter (1966)

*Grade:* 9. *Concept domain:* Biology. *Method:* 1) inductive laboratory in which the subjects designed their own experiments to solve problems, 2) inductive demonstration experiment conducted by teachers as a demonstration, 3) deductive laboratory-planned experiment following a lecture. *Measures:* Lorge-Thorndike Intelligence Test; Watson-Glaser Critical Thinking Appraisal; pre-post test by investigator. *Conclusions:* No differences among inductive and deductive methods for the teaching of facts, application of principles, for utilization of laboratory techniques; "There was some indication" (no data provided) that the inductive approach was more effective in teaching cause and effect relationships, making judgments on evidence, and evaluation of arguments.

Menefee (1965)

*Grades:* 1 to 6. *Concept domain:* Mass and volume. *Method:* "Decision-making situations," differentiation of mass or volume of pairs of objects; use of science instruments for gathering and using evidence. *Measures:* Amounts wagered by student for confidence in decision, measures of IQ, confidence, and manipulation of materials. *Conclusions:* No relation of IQ and confidence in "decision-making"; confidence related to age; decreasing errors for older children, some increase in performance for 5th and 6th grades.

Scheffer (1965)

*Grades:* College. *Concept domain:* Genetics. *Method:* The experimental group taught by inductive laboratory method, the control group taught by a "traditional method." *Measures:* TOUS; semantic differential type attitude measure and achievement test by investigator; Kuder Preference Record (Science subscale). *Conclusions:* No difference between inductive lab group and traditionally taught group.

Tanner (1968)

*Grade:* 9. *Concept domain:* Mechanics of simple machines. *Method:* Self-instructive program; three treatments—1) expository-deductive, 2) discovery-inductive, and 3) unsequenced-discovery. *Measures:* Comprehension, retention, transfer and interest made by investigator. *Conclusions:* No significant differences by treatment groups; comprehension and retention higher for experimental group, but no differences in transfer.

### GUIDED DISCOVERY

Lansdown and Dietz (1965)

*Grades:* Elementary teachers in workshop. *Concept domain:* Effects of light on colored paper. *Method:* Free experimentation vs. guided experimentation; four treatments—1) Read text then experiment, 2) experiment then read text, 3) pretest, then experiment, and 4) pretest, experiment, then read. *Measures:* Tests by authors used as pre-post test according to experimental design. *Conclusions:* Use of the pretest guided students to discover particular facts; more discoveries in free experimental without the pretest or textbook reading; reading of text led to some invalid discoveries; the students given the pretest were less interested in the experiments.

Rainey (1965)

*Grade:* 12. *Concept domain:* Chemistry. *Method:* Chemistry lab exercises with detailed instruction vs. same exercises in forms or problems. *Measure:* Standardized High School Chemistry Test of the American Chemical Society; Cooperative Chemistry Test; and Laboratory and Performance Tests by the investigator. *Conclusions:* No differences in achievement on any of the chemistry measures; non-directed group scored higher on the performance test, subjective comments in reference to group scored higher on the performance test, 1) produced better write-up of experiments, 2) took longer to adjust to lab than the directed group, 3) at the start the non-directed group voiced objections to the method, but later took pride in being able to carry through an experiment on their own, and 4) spent extra time checking results.

Salstrom (1966)

*Grade:* 6. *Concept domain:* Causality and explanation. *Method:* Oral lessons vs.



independent work from a battery of hypotheses-questions. *Measures*: California Test of Mental Maturity; STEP (Science). *Conclusion*: Greater conceptual gains for guided discovery.

Thomas (1968)

*Grade*: 8. *Concept domain*: Earth Science. *Method*: Didactic method vs. guided discovery. *Measures*: TOUS; Watson-Glaser Critical Thinking Appraisal; TAB Inventory of Science Processes. *Conclusions*: For TOUS the only gains were for guided discovery; the didactic method was superior for factual-conceptual achievement.

Two brief reports in *Dissertation Abstracts*, Gentry (1965) and Hamilton (1966), seem to indicate no significant advantage in those discovery-inductive methods in which single-concept film loops are used. Kellogg (1966) reported a study on the teaching of analysis and interpretation of graphs to teacher-preparation college students. The use of laboratory-discovery vs. demonstration-discussion resulted in no measured differences. A study by N. Scott (1966), comparing inquiry teaching to traditional teaching in grades 4-6, appears to have resulted in a few differences in student performance. Unfortunately, the findings of this study may be mostly the result of random variation due to the methods of analysis employed.

As I indicated in the introduction to this section, the enormous range of meanings assigned to "discovery" renders generalizations almost impossible on the bases of the above studies. In addition to the critiques in Shulman and Keislar's *Learning by Discovery* (1966), a theoretical study led Stokes (1963) to conclude that lack of evidence made it impossible to support the claims made for discovery learning. Newton (1968) made a number of critical comments on inquiry teaching. Inquiry teaching was also used by Sweeters (1968) in a science program for educable mentally retarded children. No significant differences were found on a variety of measures. In general, it seems reasonable to conclude that the various claims made for inquiry or discovery learning remain unsubstantiated, at least in the science education literature identified here.

### *Creative and Critical Thinking*

Studies in creative thinking are generally related to Torrance's work. However, the relation among the studies listed below seems to be due more to the use of the Torrance Tests than to attempts to contribute to the theory. Studies in critical thinking are marked more by the use of the Watson-Glaser Critical Thinking Appraisal than by their relevance to any conceptualization of critical thought.

### CREATIVE THINKING

Clark (1968)

*Grade*: 8. *Method*: Use of Inquiry Training by Science Research Associates vs. traditional teachings program. *Measures*: Torrance Creative Thinking Test; subject matter test. *Conclusions*: Traditional class superior to inquiry teaching for subject-matter achievement and creative thinking ability; girls scored higher on creative thinking tests than boys.

DeRoche (1966)

*Grade: 6. Method:* Creative exercises vs. traditional exercises in space achievement and creative thinking; two groups—1) creative exercise and brainstorming sessions, 2) traditional teaching. *Measures:* Lorge-Thorndike Intelligence Test; Minnesota Tests of Creative Thinking; Space Science Achievement test by investigator. *Conclusions:* Significant relationship between IQ-creative thinking and IQ-achievement. The abler students in experimental group superior on posttest of fluency, flexibility, and elaboration to the control group; females superior in fluency.

Tating (1965)

*Grade: 4. Method:* Creative thinking considered as an "operant behavior," defined as "divergent"; training on science words as stimuli; the children were encouraged to give original responses on a posttest. *Measures:* Pre-posttest by investigator; scored for diversity and originality. *Conclusion:* Training group produced more divergent responses; more divergent responses for the group trained on science words than for the group trained on non-science words.

## CRITICAL THINKING

George (1965)

*Grades:* High School. *Method:* Comparison of BSCS students and those in traditional biology on critical thinking. *Measures:* Watson-Glaser; Otis Quick Scoring Mental Ability test. *Conclusions:* Blue Version students scored significantly better on critical thinking test controlling for IQ; no differences for Green and Yellow Versions of BSCS and traditional biology course.

Henkel (1967)

*Grades:* College students. *Method:* PSSC students vs. students in traditional physics course. *Measures:* Watson-Glaser; Iowa placement Exams and Physics test; *Measures* of science background and mathematics background. *Conclusions:* Only PSSC students had significant gains in critical thinking; greater growth for students with previous science courses.

Kastrinos (1964)

*Grades:* High school advanced biology. *Methods:* Two methods of teaching—text-book-recitation group vs. principles-critical thinking group. *Measures:* Diagnostic Reading test; Lorge-Thorndike Intelligence Test; University of Illinois Test of Ability to Judge Interpretation of Data; Problem-solving Test; Nelson Biology test; Watson-Glaser; Kastrinos Critical Thinking Test. *Conclusions:* No difference in groups for gain of factual information or interpretation of data; experimental group higher on critical thinking test by investigator.

Rickert (1967)

*Grade:* first year college. *Method:* Special course designed to provide opportunities to analyze problems, examine assumptions, collect and organize data, and test hypotheses in physical sciences; compared to two other physical science courses. *Measures:* A.C.E. Test of Critical Thinking and STEP tests. *Conclusions:* Experimental course superior in developing "ability to think critically" for all ability groups.

As in many other studies of creative thinking, sex differences were indicated in the work of Clark and DeRoche. The studies by Tating, Clark and DeRoche differ on so many dimensions that no generalized conclusions seem possible. The critical thinking research studies listed show that performance on critical thinking tests can be enhanced when special teaching methods or curricula are utilized. This fact should not be taken lightly since it is one of the few consistent findings in studies reviewed in this chapter.

## Other Learning Studies

In this section, several other types of studies are discussed; they are grouped together simply as a result of the particular categorization scheme employed in this paper. The grouping of studies in an "other" category

should not be interpreted to mean that these studies are any less worthy. In fact, the studies identified in this section were perhaps among the best conducted during the period reviewed.

A study by Rowe (1965) is remarkable not only for its ingenious methods, but because it is one of the rare examples of the careful analysis and use of the term "concept." Rowe pointed out that an object may be assigned to different conceptual sets at different times. What is frequently critical is the *context* in which an object is embedded. Consequently, Rowe was led to formulate a mode of learning which she termed *context learning* in contrast to *concept learning*. She found that some modes of processing data by subjects in the experiment were more efficient than others. Since so many new ideas are presented, the original paper should be consulted.

Another category of science education study that has recently become more frequent is *cognitive preference*. Atwood (1967) compared students in CHEM study on four dimensions of preference: 1) memory of facts or terms, 2) fundamental principles, 3) critical questioning of information, and 4) practical application. Using a cognitive preference examination devised by the investigator and two achievement tests, Atwood found that "students who showed a strong preference for memory of specific facts were at a disadvantage in CHEM study relative to those who showed a strong preference for fundamental principles, critical questioning of information, and practical application. The data suggest a strong preference for practical application in combination with a preference for fundamental principles or critical questioning may be most advantageous in this respect."

Peterson and Lowery (1968) investigated curiosity factors in a group of first grade pupils. Children were brought individually to a "waiting room" containing a variety of science materials. The children were told that they would be interviewed by someone and that they could do as they wished while waiting for the interviewer. The children became involved with the science materials and an observer then noted the time and intensity of the involvement; the observer used special measurement methods devised for the experiment. The techniques used in this study provide a setting that is very promising for future work in curiosity and other areas of study. Curiosity seems to be a fertile area for research; hopefully, additional work will be done in this domain. An excellent review of research and discussion papers on curiosity and science education was prepared by Cunningham (1966b).

In his study on frustration, Ho (1966) employed some unusual methodological techniques, e.g., heart rate as a measure of frustration. Although the task involved the learning of geometrical shapes, the techniques could be adapted to science tasks.

Another study noteworthy for its techniques and domain of investigation is Dyasi's (1966) study of affective behaviors associated with the learning of science. Although the cognitive domain received considerable attention in recent years, work in the affective domain is quite rare in science education. Dyasi used observation, interviews and written instruments to secure data on twenty gifted students. The techniques used in this study could be employed to investigate different types of children, categorized by affective criteria, in relation to science learning.

The studies grouped as "other learning studies" seem particularly useful for their research methods and their domain of investigation. The more traditional categories (concept learning, problem solving, etc.) while far from exhausted as research areas, do seem to suffer from a number of methodological and conceptual difficulties. Mere newness or novelty, however, is surely not a sufficient criterion for rejuvenating learning studies in science education. High on a list of guidelines for future work might be the realization that many of the studies reviewed here are too simplistic for the nature of the phenomena investigated. Perhaps a return to fundamental psychological and philosophical issues would be more promising than continuing assaults on problems using moribund techniques. The old pathways seem to have been so overtraveled that the ground is rather barren.

### *Theoretical Perspectives on Learning*

During the past decade, sufficient curriculum products have been made available that one can assess what the development craft has produced. Curriculum development shares with other crafts a reliance on the theoretical, i.e., conceptual schema invented to bring the existential world into temporary order. By examining the curricula and the procedures used in their development, it seems obvious that curriculum craftsmen have turned to quite diverse sets of abstractions to guide their work, in fact, to make getting on with their work possible.

Interestingly, after more than a decade of new curriculum development there is suddenly an increase in the number of papers in the science education literature devoted to examinations of various psychological ideologies from which researchers have taken advice. There is considerable danger that the attempts to point out limitations of a particular ideology will automatically lead their proponents to identify critiques as heretical attacks. This could further entrench positions and foment divisive discourses. The field of course development generally is replete with ideological stances, but science education has been particularly noteworthy for adopting different schools of learning psychology.



## Bruner

In the papers examined for the review, the work most often cited was Bruner's *The Process of Education* (1960). Authors most frequently referenced the second chapter on the concept of structure. The polarized reactions in the broader educational literature to Bruner's concepts of structure were surprising; either his hypothetical statements were accepted as establishment principles appropriate for use as curricula guidelines or critics screamed that the evidence for the hypothetical statements was not yet in. Rereading Bruner's chapter on structure renders these reactions all the more puzzling since Bruner continually repeats such phrases as "... something on which a great deal of research is needed . . .," "research is in progress to elucidate the matter . . .," "... claims in need of detailed study," etc. Another reception of the Bruner proposals has been to weave Brunerian themes into discussions of science and the teaching of science (see Carlson, 1967c; Fischler, 1965; and Ivany, 1966).

## Gagné

Gagné presented a different psychological view of learning which became embodied in the AAAS Elementary Science Program. A detailed exposition was presented in Gagné's 1965 book *The Conditions of Learning*. H. Scott (1966) compared the behaviors listed in the AAAS Science Program with Bloom's *Taxonomy of Educational Objectives: The Cognitive Domain*.

Gagné's views are very similar to B. F. Skinner's general position on emphasizing shaping the behavior of an organism. Gagné's arguments for strategies of curriculum development parallel the Skinnerian arguments for programed instruction. The emphasis on clear-cut behavioral goals in the behavioristic ideology lends itself, by extension, to the fundamental ideas of systems analysis and all its subspecies (see Haraseyke and Fanning, 1969). The literature on systems analysis and its applications to business, government, economics, the military, sociology, the social and behavioral sciences and education has grown so enormous that it constitutes a specialized branch of study. The major concern for this review is that the melded behavioristic-systems analysis ideology imposes a particular model of man compatible with its tenets, while other models of man are likely to be dismissed as obscure, fuzzy or romantic. In the systems analysis ideology "learning" is reduced to the external manifestations of modified behavior, and the model of "learning man" emphasizes, on the input side, the ordering of the external environment. It is fascinating to trace these tenets to their philosophical-historical contexts, and such an analysis would be a welcome addition to science education literature. Numerous controversies could arise between behaviorist's ideologies and those with alternative views.

Current trends in science learning and in education generally seem to indicate that the analyses and predictions which the French sociologist Jacques Ellul (*The Technological Society*) made in 1954 are increasingly being borne out. Ellul argued that the various spheres of man's endeavors are so rich in techniques (means) that men are quickly reaching the stage where what were formerly treated as means are now being treated as ends to which they will have to adapt. The behavioristic framework is a means-ridden ideology that provides techniques for learning, curriculum construction and evaluation. There are very real advantages to the behavioristic framework; it suggests powerful techniques of control and it provides clarity and feedback possibilities. However, it would be extremely ironic if, as Ellul predicted, the final educational outcome were to be an immense control system which educator and student alike must serve. Instead of reiterating the behavioristic-systems analysis tenets in response to justified cautions, educators would do better to read Ellul's book and search for alternatives to his pessimistic vision, however "nebulous and obscure" these alternatives might be.

*Additional References:* Esler (1968); Kolb (1967); O'Toole (1968a, 1968b).

### Ausubel

Increased attention is being given to Ausubel's theory of meaningful, verbal reception learning and retention. Ausubel discussed his views in *The Psychology of Meaningful Verbal Learning* (1963) and in *Educational Psychology: A Cognitive View* (1968). Several critiques of contemporary science curricula were presented by Ausubel (1965, 1966) from this point of view. Ausubel's theory of "meaningful verbal learning" is eclectic, drawing on many sources for its structure. Ausubel's theory appears to be a very promising framework for science learning research in those conceptual aspects of science which can be represented by the symbolic (words, signs, symbols). For researchers whose epistemology of science leans toward symbolic representation, Ausubel's theories are rich with possibilities.

*Additional References:* Novak (1965, 1966); Schulz (1966).

### Piaget

The name of Jean Piaget looms large during the period under review. In 1964 most of Issue 3 of the *Journal of Research in Science Teaching* was devoted to papers concerned with selected aspects of Piaget's work. Piaget's paper "Development and Learning" was included. Piaget's work has suffered an unusual fate in the English-speaking world. In the 1930's the descriptive phenomena of children's behavior given in Piaget's first

five books were brutally criticized by American and British authors, the phenomena described were usually explained away either as artifacts of the "clinical method" or as problems of language, or both. After twenty years of endless replication studies by American and British researchers and a great deal of additional work by the Genevans, the existence of the phenomena was finally accepted. At present, however, there is the strong tendency to account for the Piagetian findings by explanatory frameworks embedded within the empirical-positivistic tradition of the Anglo-Saxon world.

Incorporating Piaget's ideas into our own traditions is quite natural, but it severely hinders understanding Piaget on his own terms and within his tradition—especially in the discussions contained in the first and last chapters of any Piaget book. Readers are usually fascinated by the middle chapters in which Piaget describes all those cute things children do or do not do. Unfortunately, readers usually neglect the philosophical problems that Piaget poses in the early chapters and also seem to neglect the concluding chapter describing whatever light has been cast on the philosophical problems resulting from the empirical investigations. Widely interpreted and summarized as a developmental psychologist, Piaget has been rediscovered as an epistemologist operating within the framework of biology rather than within the positivistic-mechanistic ideology. David Elkin's introduction to Piaget's *Six Psychological Studies* is an excellent statement that clarifies a number of misunderstandings and clearly emphasizes Piaget's fundamental epistemological concerns.

It is most unfortunate for science educators that Piaget's books dealing explicitly with epistemology, particularly those dealing with the biological and physical sciences, have not been translated. It is also regrettable that of the 24 volumes in the genetic epistemology monograph series only one is available in English. In 1969 Hans Furth's outstanding book, *Piaget and Knowledge*, was published. Furth's book may well mark a turning point in establishing Piaget's biological framework as the fundamental domain for understanding not only his theory of intelligence but also his theory of knowledge in general. The book should clarify many severe misunderstandings and hopefully will induce many readers to return to Piaget's writings with a greater appreciation of the fundamental differences between his work and the Piaget-inspired American research.

In the science education literature that I reviewed, frequent reference was made to the importance of giving children opportunities to act on and transform materials as a pedagogical strategy consistent with Piaget's views of development. Articles by Fischler (1968) and Carlson (1967a) contain discussions on science curriculum and teaching strategies in relation to selected aspects of the Geneva School. A study on the acquisition

of the conservation of substance by Carlson (1967b) investigated the employment of two levels of verbal instruction, high and minimal, and demonstration vs. direct manipulation by the child. All treatments improved the performance of the children on a criterion task compared to a control group. The Carlson study illustrates an area of Piagetian studies that has been of considerable interest to American researchers, namely the acceleration of the acquisition of a particular scheme. The whole issue about whether or not the acquisition of schema can be accelerated has led to needless controversy and is a good example of the kind of issue Furth's book may help clarify.

Raven's (1967) excellent study of the concept of momentum in children contributed additional information to the Piagetian corpus of research; it is profoundly different from the concept studies reported in the first half of this review. Raven's colorful analysis of the sequence of concepts necessary for understanding momentum illustrates the kind of research so badly needed to make Piaget's very general description of concrete operations useful in curriculum development. There has been speculation that the Piaget development sequence would be useful in devising curriculum sequences for elementary school science. This is probably true, but Piaget described the stage of concrete operations only in very general characteristics of children seven to twelve years old. Obviously, the specific characteristics of a seven year old differ extensively from those of a twelve year old. Piaget is concerned with describing the "generalized knower" (the epistemic subject), and while such description sheds profound insight on epistemological problems, it lacks the fine-structure required to be truly helpful in detailed curriculum specifications and prescriptions. It seems that American researchers, who have particular talents in the techniques of empirics, could make a significant contribution to Piagetian studies by investigating the fine-structure and reporting what happens in detail within the stage of concrete operations between seven and twelve years of age for very specific science tasks. Raven's study of momentum is a very fine illustration of the kind of work needed.

In Geneva, Inhelder and Sinclair are currently attacking the problem of schema interaction in a series of learning studies. In Montreal, Laurendeau and Pinard have published a detailed study of schema interaction in the child's early concepts of space. The American style of research also appears to be very suitable for investigating this problem. Piaget has long maintained that schema are interrelated, never operating in isolation but always within the matrix of a structured whole. More knowledge about the interaction of schema is necessary before Piagetian theory can be really useful in learning studies or developing curriculum sequences. Fine-structure and schema interaction are such intriguing areas of research that it would seem far more profitable to tackle specific problems within



these two areas than to generate yet another replication variation. Shedding even a tiny glimmer of light on either of these problems by a doctoral candidate would be a far better test of his research competency than a virtuoso technical dissertation on a trivial matter.

The ultimate publication of the current work on causality by Piaget and his colleagues will be of great interest to science educators. The results of a five-year research program devoted to causality may well provide information of immense value to science learning, science curricula and future research.

### Bibliography

- Adler, Leona K. The Development of Concepts of Space, Matter and Energy in Students at the College Level. *Journal of Research in Science Teaching* 4: 32-41; 1966.
- Anderson, Hans O. An Analysis of a Method for Improving Problem-Solving Skills Possessed by College Students Preparing to Pursue Science Teaching as a Profession. Doctor's thesis. Bloomington: Ind. Univ., 1966. Abstract: *Dissertation Abstracts* 27: 3332A; No. 10, 1967.
- Anderson, Hans O. Problem-Solving and Science Teaching. *School Science and Mathematics* 57: 243-51; 1967.
- Anderson, Hans and Weigand, James. Instructional Theory, Problem-Solving and Science Teaching. *School Science and Mathematics* 67: 483-90; 1967.
- Anderson, Ronald D. Children's Ability to Formulate Mental Models to Explain Natural Phenomena. *Journal of Research in Science Teaching* 3: 326-32; 1965.
- Atwood, Ronald K. CHEM Study Achievement Among Groups Classified by Cognitive Preference Scores. *Journal of Research in Science Teaching* 5: 154-59; 1967.
- Ausubel, David P. An Evaluation of the "Conceptual Schemes" approach to Science Curriculum Development. *Journal of Research in Science Teaching* 3: 254-64; 1965.
- Ausubel, David P. An Evaluation of the BSCS Approach to High School Biology. *American Biology Teacher* 28: 176-86; 1966.
- Aylesworth, Thomas G. The Need for Problem-Solving. *Science Education* 49: 156-62; 1965.
- Bruner, Jerome S. *The Process of Education*. Cambridge, Mass.: Harvard Univ. Press, 1960.
- Brusini, Joseph A. An Experimental Study of the Development of Science Continua Concepts in Upper Elementary and Junior High School Children. Doctor's thesis. Evanston, Ill.: Northwestern Univ., 1966. Abstract: *Dissertation Abstracts* 27: 2094A; No. 7, 1967.
- Butts, David P. The Relationship of Problem-Solving Ability and Science Knowledge. *Science Education* 49: 138-46; 1965.
- Butts, David and Jones, Howard. Inquiry Training and Problem Solving in Elementary School Children. *Journal of Research in Science Teaching* 4: 21-27; 1966.
- Butts, David and Jones, Howard. The Development of the TAB Science Test. *Science Education* 51: 463-73; 1967.
- Carlson, Jerry. Effects of Instruction on the Concept of Conservation of Substance. *Science Education* 51: 138-45; 1967a.
- Carlson, Jerry. Developmental Psychology and its implications for Science Education. *Science Education* 51: 246-50; 1967b.
- Carlson, Jerry. Science and the Curriculum. *Science Education* 51: 251-54; 1967c.
- Clark, Billy M. An Experiment in Cultivating Creative Thinking Abilities in the Classroom. Doctor's thesis. Ames: Ia. State Univ., 1968. Abstract: *Dissertation Abstracts* 29: 1158A; No. 4, 1968.
- Coulter, John C. The Effectiveness of Inductive Laboratory, Inductive Demonstration, and Deductive Laboratory in Biology. *Journal of Research in Science Teaching* 4: 185-86; 1966.

- Cunningham, John D. Rigidity in Children's Problem Solving. *School Science and Mathematics* 66: 377-89; 1966a.
- Cunningham, John D. On Curiosity and Science Education. *School Science and Mathematics* 66: 805-15; 1966b.
- Dallas, Lee A. A Comparison of the Effects of Two Structures of Knowledge on the Application of Science Concepts by Preservice Elementary School Teachers. Doctor's thesis. Philadelphia: Temple Univ., 1968. Abstract: *Dissertation Abstracts* 29: 1804A; No. 6, 1968.
- Dennis, David M. The Introduction of Concepts in Kinetic Molecular Theory to Children. *Journal of Research in Science Teaching* 4: 106-11; 1966.
- DeRoche, Edward F. A Study of the Effectiveness of Selected Creative Exercises on Creative Thinking and the Mastery of a Unit in Elementary Science. Doctor's thesis. Storrs: Univ. of Conn., 1966. Abstract: *Dissertation Abstracts* 27: 4162A; No. 8, 1966.
- Dyasi, Hubert. An Exploratory Investigation of Certain Affective Behaviors Associated with the Learning of Science. Doctor's thesis. Urbana: Univ. of Ill., 1966. Abstract: *Dissertation Abstracts* 27: 3770A; No. 10, 1967.
- Fischler, Abraham. Science, Process, the Learner: A Synthesis. *Science Education* 49: 402-409; 1965.
- Fischler, Abraham. Implications of Structure for Elementary Science. *Science Education* 52: 277-84; 1968.
- Gentry, Castelle G. Relative Effectiveness of Discovery and Expository Methods of Teaching Concepts through the Single-Concept Film. Doctor's thesis. East Lansing: Mich. State Univ., 1965. Abstract: *Dissertation Abstracts* 26: 4497; No. 8, 1966.
- George, Kenneth D. The Effect of BSCS and Conventional Biology on Critical Thinking. *Journal of Research in Science Teaching* 3: 293-99; 1965.
- Grant, Dorothy A. Problem Solving in Teaching with Emphasis on Instruction in Science Programs for the Early Adolescent. Doctor's thesis. New York: Columbia Univ., 1965. Abstract: *Dissertation Abstracts* 26: 5912; No. 10, 1966.
- Hamilton, Donald K. The Relative Merits of Two Teaching Methods in Improving the Skills of Observation and Inference-Making of Seventh and Eighth Grade Science Students. Doctor's thesis. Bloomington: Ind. Univ., 1966. Abstract: *Dissertation Abstracts* 27: 3340A; No. 10, 1967.
- Haraseyke, Harry and Fanning, John. Information System Science—Challenge to Education. *The Science Teacher* 36: 31-35; Jan. 1969.
- Harris, William and Lee, Verlin. Mental Age and Science Concepts: A Pilot Study. *Journal of Research in Science Teaching* 4: 275-88; 1966.
- Henkel, E. Thomas. The Effects of Instruction in Physics Upon the Critical Thinking Ability of Undergraduate Students. *Journal of Research in Science Teaching* 5: 89-94; 1967.
- Ho, C. J. The Effects of Frustration on Intellectual Performance. *Science Education* 50: 457-60; 1966.
- Inventasch, Harvey. A Comparison of the Effects of Teacher-Directed and Self-Directed Problem Solving on Attitudes and Understandings in Science. Doctor's thesis. New York: N. Y. Univ., 1968. Abstract: *Dissertation Abstracts* 29: 497A; No. 2, 1968.
- Ivany, J. W. George. Psychological Aspects of Structure in Science. *The Science Teacher* 33: 36, 37; May 1966.
- Kantz, Paul T. An Investigation of Elementary School Children's Concepts of Selected Terms and Principles in Science in 1950 and 1964. Doctor's thesis. Austin: Univ. of Tex., 1968. Abstract: *Dissertation Abstracts* 29: 1680A; No. 6, 1968.
- Kastrinos, William. The Relationship of Two Methods of Teaching to the Development of Critical Thinking by High School Students in Advanced Biology. *Science Education* 48: 187-95; 1964.
- Kellogg, Maurice G. The Effect of Laboratory-Discovery Methods and Demonstration-Discussion Methods upon Elementary Science Methods Students' Abilities to Analyze and Interpret Graphs. Doctor's thesis. Bloomington: Ind. Univ., 1966. Abstract: *Dissertation Abstracts* 27: 345A; No. 2, 1966.
- Kolb, John R. Effects of Relating Mathematics to Science Instruction on the Acquisition

- of Quantitative Science Behaviors. *Journal of Research in Science Teaching* 5: 175-82; 1967.
- Lansdown, Brenda and Dietz, Thomas. Free Versus Guided Experimentation. *Science Education* 49: 175-82; 1965.
- Mahan, Luther A. What is the Problem Solving Method of Teaching? *Science Education* 51: 332-43; 1967.
- Menefee, Robert W. *Measuring Elementary School Children's Ability to Use Evidence from Scientific Instruments in Decision-Making Situations*. Doctor's thesis. Columbus: Ohio State Univ., 1965. Abstract: *Dissertation Abstracts* 27: 117A; No. 1, 1966.
- Newton, David E. The Dishonesty of Inquiry Teaching. *School Science and Mathematics* 68: 807-10; 1968.
- Novak, J. D. A Model for the Interpretation and Analysis of Concept Formation. *Journal of Research in Science Teaching* 3: 72-83; 1965.
- O'Toole, Raymond J. *A Study to Determine Whether Fifth Grade Children Can Learn Certain Selected Problem Solving Abilities through Individualized Instruction*. Doctor's thesis. Greeley: Colo. State College, 1966. Abstract: *Dissertation Abstracts* 27: 3781A; No. 11, 1967.
- Pella, Milton and Zeigler, Robert. Use of Mechanical Models in Teaching Theoretical Concepts. *Journal of Research in Science Teaching* 5: 138-50; 1967.
- Peterson, Rita and Lowery, Lawrence. A Study of Curiosity Factors in First Grade Children. *Science Education* 52: 347-52; 1968.
- Rainey, Robert G. The Effects of Directed Versus Non-Directed Laboratory Work on High School Chemistry Achievement. *Journal of Research in Science Teaching* 3: 286-92; 1965.
- Raven, Ronald. The Development of the Concept of Momentum in Primary School Children. *Journal of Research in Science Teaching* 5: 216-23; 1967.
- Rickert, Russell K. Developing Critical Thinking. *Science Education* 51: 24-27; 1967.
- Roberts, John B. *A Study of the Problem Solving Processes of Successful and Non-Successful Problem Solvers in Ninth Grade Science*. Doctor's thesis. Tuscaloosa: Univ. of Ala., 1964. Abstract: *Dissertation Abstracts* 25: 7088; No. 12, 1965.
- Rowe, Mary Budd. Influence of Context-Learning on Solution of Task-Oriented Science Problems. *Journal of Research in Science Teaching* 3: 12-18; 1965.
- Salstrom, David. *A Comparison of Conceptualization in Two Types of Guided Discovery Science Lessons*. Doctor's thesis. Kent, Ohio: Kent State Univ., 1966. Abstract: *Dissertation Abstracts* 28: 407A; No. 2, 1967.
- Scheffler, William. A Comparison Between Inductive and Illustrative Laboratories in College Biology. *Journal of Research in Science Teaching* 3: 218-23; 1965.
- Schulz, Richard W. *The Role of Cognitive Organizers in the Facilitation of Concept Learning in Elementary School Science*. Doctor's thesis. West Lafayette, Ind.: Purdue Univ., 1966. Abstract: *Dissertation Abstracts* 27: 3784; No. 11, 1967.
- Scott, Harry. *Cognitive Analysis of a Curriculum; An Application of Taxonomy of Educational Objectives Handbook I; Cognitive Domain to Science—A Process Approach*. Doctor's thesis. New York: Columbia Univ., 1966. Abstract: *Dissertation Abstracts* 27: 3637; No. 11, 1967.
- Scott, Norval C. The Strategy of Inquiry and Styles of Categorization. *Journal of Research in Science Teaching* 4: 143-53; 1966.
- Shulman, Lee S. and Keislar, Evan R. (editors). *Learning by Discovery*. Chicago: Rand McNally, 1966.
- Stokes, William W. *An Analysis and Evaluation of Current Efforts to Improve the Curriculum by Emphasis on Disciplinary Structure and Learning by Discovery*. Doctor's thesis. Gainesville: Univ. of Fla., 1963. Abstract: *Dissertation Abstracts* 25: 329; No. 1, 1964.
- Sweeters, William G. *Discovery Oriented Instruction in Science Skills for Educable Mentally Retarded Children*. Doctor's thesis. Greeley: Colo. State College, 1968. Abstract: *Dissertation Abstracts* 29: 1376A; No. 5, 1968.
- Tanner, Richard T. *Expository-Deductive vs. Discovery-Inductive Programming of Physical Science Principles*. Doctor's thesis. Stanford: Stanford Univ., 1968. Abstract: *Dissertation Abstracts* 29: 1480A; No. 5, 1968.
- Tating, Marcela T. *Priming Creative Thinking in Elementary School Science*. Doctor's



- thesis. Los Angeles: Univ. of Calif., 1965. Abstract: *Dissertation Abstracts* 26: 3177; No. 6, 1966.
- Their, Herbert D. A Look at First Grader's Understanding of Matter. *Journal of Research in Science Teaching* 3: 84-89; 1965.
- Thomas, Barbara S. *An Analysis of the Effects of Instructional Methods upon Selected Outcomes of Instruction in an Interdisciplinary Science Unit*. Doctor's thesis. Iowa City: Univ. of Ia., 1968. Abstract: *Dissertation Abstracts* 29: 1830A; No. 6, 1968.
- Trent, John. The Attainment of the Concept "Understanding Science" Using Contrasting Physics Courses. *Journal of Research in Science Teaching* 3: 224-29; 1965.
- Wallace, J. G. *Concept Growth and the Education of the Child*. National Foundation for Educational Research in England and Wales, Occasional Publication Series No. 12. The foundation (The Mere, Upton Park, Slough, Bucks., England), 1965.

### Additional References

- Esler, William. An Exercise in the Development of Process Skills. *School Science and Mathematics* 68: 821-23, 1968.
- Kolb, John R. Effects of Relating Mathematics to Science Instruction on the Acquisition of Quantitative Science Behaviors. *Journal of Research in Science Teaching* 5: 175-82; 1967.
- Novak, J. D. A Model for the Interpretation and Analysis of Concept Formation. *Journal of Research in Science Teaching* 3: 72-83; 1965.
- Novak, Joseph. The Advance of Science Education. *Education* 87: 223-27; 1966.
- O'Toole, Raymond J. Individualized Elementary School Science. *Science Education* 52: 376-80; 1968a.
- O'Toole, Raymond J. The Effectiveness of Individualized Elementary School Science. *Science Education* 52: 381-84; 1968b.
- Pella, Milton and Carey, Russell. Levels of Maturity and Levels of Understanding for Selected Concepts of the Particle Nature of Matter. *Journal of Research in Science Teaching* 5: 202-15; 1967.
- Pella, Milton and Strauss, Nyles. The Relationship between Concept Attainment and Level of Maturity. *Journal of Research in Science Teaching* 5: 299-310; 1967.
- Pella, Milton and Triesenberg, Henry. Three Levels of Abstraction of the Concept Equilibrium and its Use as an Advance Organizer. *Journal of Research in Science Teaching* 6: 11-21; 1969.
- Pella, Milton and Voelker, Alan. Teaching the Concepts of Physical and Chemical Change to Elementary School Children. *Journal of Research in Science Teaching* 5: 311-23; 1967.
- Schaal, William L. Scientific Concepts in the Junior High School Mathematics Curriculum. *School Science and Mathematics* 65: 614-25; 1965.

### AUTHOR

MAURICE BELANGER Address: Harvard University, Cambridge, Massachusetts Title: Lecturer to Education Age: 40 Degrees: B.S. and M.A., Univ. of New Hampshire; Ed.D., Harvard Univ. Specialization: Curriculum and Supervision; Elementary School Science; Child Development.



# THE VOCATIONAL GUIDANCE QUARTERLY

*Official Publication of the National Vocational Guidance Association*

- Concerns the role of work in the life of man.
- Reports vocational guidance research.
- Covers all aspects of vocational counseling in schools, colleges, government, business, and industry as well as topics of interest to labor and management.
- Shows how vocational and educational goals mesh.
- Publishes latest information on career selection and outlook in today's world.
- Presents viewpoint of potential employers.
- Reviews new books and current career films.
- Classifies and codes current occupational materials.

## ARTICLES ACCEPTED FOR PUBLICATION INCLUDE:

*Vocationalism in Higher Education*

*Assessing Work Potential of the Handicapped in the Public Schools*

*The Concept of Work Values*

*Current Status of the Measurement of Vocational Maturity*

- Four issues, September, December, March, June—\$5.
- Enclose payment with your order (subscriptions begin when payment is received) and mail to:

## AMERICAN PERSONNEL AND GUIDANCE ASSOCIATION

Subscription Department E-12

1607 New Hampshire Avenue, N.W.

Washington, D.C. 20009

## 2: CURRICULUM DEVELOPMENT AND INSTRUCTIONAL MATERIALS

Herbert A. Smith

*Colorado State University*

Limitations of time and space have made it almost impossible to produce an accurate, impartial and comprehensive review of research on the curriculum development and instructional materials in science education which have been produced and disseminated at an immensely increasing rate since the last review. Hence this chapter is highly selective; important studies and some large areas of potential interest have been omitted.

The ERIC Center for Science Education (1968a, 1968b) provided a number of invaluable bibliographies. Lockard (1968) compiled a series of annual reports which are extensive compendia of information on both national and international science and mathematics curricular developments. O'Hearn (1967) wrote a useful introduction to the major curriculum developments in science.

*Additional References:* Mahler (1967); National Science Teacher Association (1968).

### *Philosophical and Historical Bases of Curriculum Development*

Science education, especially in curriculum, is rent by dichotomies. Many scientists and science educators have emphasized the structure and the processes of science as a discipline and as an intellectual accomplishment. Philosophers and an increasing number of science educators have emphasized the social implications of science as a product and as a useful way of solving practical (not merely abstract) world problems. The first view is well exemplified in Kessen's (1964, p. 4) statement: "The first task and central purpose of science education is to awaken in the child, whether or not he will become a professional scientist, a sense of joy, the excitement, and the intellectual power of science. Education in science, like education in letters and the arts will enlarge the child's appreciation of the world; it will also lead him to a better understanding of the range and limits of man's control over nature."

Cohen (1964) took a more socially oriented position; he identified 11 values and goals of science education. Citizens were to be produced

who understood the role of science in society; who could communicate with scientists; and who could apply the critical attitudes, thinking, and methods of science.

Such a distinguished scientist as Pauling (1966, p. 18) pleaded for teaching the social implications of science:

Among the problems with which we may be concerned are the pollution of the atmosphere, the pollution of water supplies, fluoridation of water and use of other public health measures, contamination of the earth with pesticides, with lead from leaded gasolines, misuse of chemicals as food additives, the location of nuclear power plants in thickly populated centers, the best use of scientific and medical knowledge to decrease the amount of human suffering caused by poverty and disease, and especially the prevention of the destruction of civilization by nuclear war.

Fox (1966) also noted the general omission in the new curricula of technology in its relation to scientific progress; he criticized curriculum planners for failing to include materials on the solution of human problems through science. Stotler (1965) and Carleton (1967) maintained an intermediate position which clearly did not exclude the consideration of social implications from the science curriculum.

If the social implications of science are to receive the attention which a rational analysis indicates they deserve, they should be included in the science curriculum. There is some evidence that they are not receiving attention in other curricula. O'Hearn and Pella (1967) reported that an analysis of textbooks used in "Problems of Democracy" courses showed that science and technology as a factor in cultural change was not considered.

*Additional Reference:* Novak (1963).

### *The Role of Scientists in Curriculum Developments*

For more than a decade, professional scientists have been almost exclusively responsible for national curriculum reforms. Their influence is evident in the emphases on the structure of science, the cognitive and abstract aspects of the discipline, the difficulty of the material, the grade level at which abstract concepts are introduced, and laboratory and inquiry experiences. Opinions of their participation and success are strongly polarized. The wisdom of some of the new emphases was questioned by a few skeptics, but the scientists' judgments were uncritically accepted for a time even, strangely enough, by most science educators.

Recently, however, the literature has reflected a growing disenchantment with new science curricula. Smith (1966) and Gatewood (1968)

charged the new curricula with being "elitist." Gatewood commented that the NSF sponsored projects have "resulted in the generation of a modern, up-to-date version of the elite science curriculum that existed during the latter part of the nineteenth century," and that the material "is essentially irrelevant to the needs of a major portion of the students whom, according to both the philosophy and laws of our nation, it should now serve" (p. 19).

Kline (1965, pp. 65, 66) denounced the role scientists have played:

The curriculums have been taken over by professional [scientists] whose aim, judged by the curriculum they have produced, is to train professionals. These reformers assume that mathematics and science are ends in themselves, that students are automatically motivated, and that the goal is to rush the education so that 17-year-olds can start writing research papers. . . .

The professors who have led the new curriculum movements have not even been wise men. Because they are the products of the narrow specialization which is characteristic of modern science education, their ignorance of the cultural significance of science may be excusable. But these men have shown a presumption and an egotism which is almost unbelievable. Most of them had never set foot in a high school or elementary school classroom and had even disdained any interest in education. When they did decide to take an interest in curriculum they assumed that education is a simple, obvious matter. Of course the professional scientists have made a fiasco of reform.

A more moderate view of the role of scientists in curriculum development was expressed by Gatewood and Obourn (1963). The writers propose that scientists of stature should be leaders in and workers for reform, along with persons from all areas of the educational community; and that no single group, no matter how distinguished, should monopolize curriculum construction. In view of the developments since Sputnik, it is questionable that scientists should lead curriculum reform; perhaps they should be relegated to the role of consultants.

Lehman (1967) wrote that the potential for stressing humanistic elements exists in new curricula, but it must be emphasized. He stated that science education should contribute substantially to 1) a concept of self-actualization, 2) a concept of alternatives and consequences, 3) a concept of mutual rights, and 4) a concept of community.

### *Learning and the Science Curriculum*

The assumptions about learning which are implicit in the curriculum



reform movement have also been challenged. Sutman (1966, p. 495) wrote that the new curricula have been "used as a cloak behind which the content of the secondary level science curricula has become more and more abstract, analytical, and theoretical"; the content of the programs "has been selected at the whim of the scientist without regard to the abilities of youngsters to deal in abstractions, and with little concept for interest getting or motivational objectives." Cronbach (1964) pointed out that the principles of pedagogy implicit in the curriculum reform movements are traditional. He noted Piaget's theories of children's developmental cognitive stages and the "false" learning which may occur when advanced concepts are presented too early. The long-term, adverse effects of early exposure to advanced concepts are unknown. Such knowledge is basic to a proper evaluation of the new curricula. Bruner (1965, p. 20) took a seemingly antithetical view in his well-known dictum that "it is possible to teach any subject to any child at any age in some form that is honest—and interesting."

### *Historical Perspectives*

In an interpretative report of research efforts in elementary and junior high school science, Smith (1963) identified the directions of the early research, some of its historical antecedents, some persistent issues and problems, and some new problems. Summerlin and Craig (1966) investigated the content of chemistry books from 1800 to the present. The decline in descriptive materials from books published as recently as 1961 to 1965 is remarkable (from 60% to 5%). Descriptive materials included such topics as the history of chemistry, industrial and commercial chemical processes, and applications common in everyday life. An equally impressive increase in the number of theoretical concepts treated has occurred.

### *Current Dimensions in Science Curriculum Construction*

#### *Behavioral Objectives*

There has been intense interest recently in defining desired learner terminal behaviors. Walbesser (1963) asserted that establishing behavioral objectives will provide both teacher and student with knowledge of the character of what is to be learned. Walbesser (1966, p. 34) stated: "... the behavioral view of curriculum development possesses advantages sufficiently unique and productive to design, development, and evaluation of instructional materials so as to make behavioral description of the objectives of instructional materials an unavoidable partner of content selection."

Montague and Butts (1968) concluded that behavioral objectives do not necessarily limit the outcomes of learning and suggested that behavior-

al objectives should broaden human perception of the learning process and should result in more varied methods of teaching and evaluation. Hungerford (1968) asked some pertinent questions about determining which terminal behaviors in science are acceptable in junior high school. He then identified seven behaviorally oriented characteristics which should typify the products of the junior high curriculum.

Not all educators have been so enamored with the potential of behavioral objectives. Atkin (1968) pointed out the importance of unplanned, unanticipated learnings. He stated that undue emphasis on behavioral objectives may result in the disappearance of worthwhile learning activities that cannot easily be identified with specific behaviors. More over, identification of a behavioral objective does not indicate anything of its value. He stressed (Atkin, 1968, p. 29) that:

... ideas are taught with the richest meaning only when they are emphasized repeatedly in appropriate and varied contexts. Many of these contexts arise in classroom situations that are unplanned, but that have powerful potential. It is detrimental to learning not to capitalize on the opportune moments for effectively teaching one idea or another. Riveting the teacher's attention to a few behavioral goals provides him with blinders that may limit his range.

Gideonse (1969) raised a broader and perhaps more fundamental question when he questioned the suitability of using the academic disciplines as the basis for choosing curricular objectives. He pointed out that interdisciplinary approaches are needed to solve human problems. Although Gideonse did not push the argument this far, it seems appropriate to ask whether problems, such as those previously alluded to by Pauling, should not determine ultimate objectives.

The current interest in behavioral objectives raises other questions. Hungerford's seven points (referred to above) aptly illustrate part of the problem. What are the implications of his views of the terminal behavior desired of students in relation to science and scientists? There must be hundreds, perhaps thousands, of behaviors which might be interpreted as evidence that the student has acquired the desired learning. However, can such behaviors be uniformly observed? Are such observations valid and can they be unambiguously interpreted? It appears that too much emphasis on behavioral objectives may mire educators in a morass of minutia and cause a reversion to the sterile "specific objectivism" that once characterized education and which has long since been abandoned.

### *Scientific Literacy*

There is an increasing concern that citizens be scientifically literate.

Pella (1965) argued that physics and chemistry courses, at three or more levels of sophistication, should be available to all students in high school (except those in special classes). The need for scientific understanding is so great that Pella recommended four years of science as a requirement for graduation. However, the specific science required would vary with students. Pella, O'Hearn and Gale (1966) analyzed conceptual knowledge as it relates to scientific literacy. The authors identified the following components as necessary to scientific literacy: 1) concepts that provide a language for communication between the lay public and the scientific community, 2) concepts that provide laymen with understanding of burgeoning scientific and technological knowledge, 3) an understanding of the historical and cultural roots of science concepts, 4) science concepts that are applicable to all education, 5) concepts which free man from supernatural explanation, and 6) study of scientific facts and concepts which bear on controversial topics and which may produce desirable changes in attitudes.

Robinson (1965) raised pertinent questions about the understandings attained by students and the nature of the science they are to understand. Identification of the pertinent aspects is essential to guide the selection and organization of elements for science curricula. The author developed a thesis that scientific inductive-deductive systems are influenced by the psychological and social climate in which they are formulated and are not necessarily accepted because of scientific criteria alone. He considered such an understanding to be essential to scientific literacy.

### *Concepts, Theories and Models*

Much attention has recently been devoted to concepts, theories and models. Some writers refer to the structure of science and the need for comprehensive, unifying themes or concepts; others see such large ideas as summarizers of a vast range of experience and knowledge and as predictors of new knowledge; and still others see such constructs as contributing to the efficiency of the learning process.

Slawson (1968) considered the appropriateness of the Thirty-First Yearbook of the NSSE, the unifying themes of the new curricula, and the NSTA *Theory Into Action* statement as guides to curriculum development. He concluded that 1) any of the three statements would be acceptable to determine the natural history and technology portions of the scientific enterprise to be included in the curriculum; 2) some unified, sequential and interdisciplinary program of conceptual schemes is required to educate scientifically literate citizens; 3) the unifying themes of the new curricula would have to be restructured to encompass all of science; 4) there are difficult problems in designing curricula based on a conceptual schemes approach; 5) both content and process goals must be considered; and 6) 7)

a conceptual schemes approach would not necessarily be a K-12 sequential program.

Karplus (1964) contended that there are key concepts which are essential for a scientifically literate person, and that these concepts are not necessarily learned through ordinary experience. One infers that such concepts are essential to the structure of science. Bruner (1965, p. 21) described a theory as providing "a terse account of what is known without the burden of details. . . . It is a canny way of keeping in mind a vast amount while thinking about very little." In a somewhat similar manner, Bronowski (1968) defined a model as a description of the way nature is organized. When experiments contradict the predictions of a model, its organization is wrong. He pointed out that no model is universal and that discrepant events will invariably occur; thus, there is an historical evolution in the development of a model.

The relationship of higher order abstraction to classroom instruction is well exemplified by the work of Atkin and Novak. In describing an ongoing project, Atkin (1963, p. 129) remarked that "a delineation of content is required that reveals a potent hierarchy of conceptual schemes, a few ideas with considerably intellectual mileage that help the learner understand in the most economical manner possible a given discipline. . . ." In describing a model for interpretation and analysis of concept formation, Novak (1965, p. 72) stated that "the *structure* of a science may be viewed as the system of major generalizations or concepts together with the *process* by which these concepts are obtained and enlarged." The purpose of Novak's study was to outline a model for interpreting the mental process of concept formation which might be useful for classroom research. This study may prove to be extremely helpful in ordering and sequencing experiences. Such a model provides a longitudinal time dimension for the instructional process which has too often been missing.

Pella and his associates reported a number of studies of learning and teaching selected science concepts. Pella and Carey (1967-1968) studied the levels of understanding of concepts about the particle nature of matter achieved by children in grades 2 to 5. They found that some concepts could be learned successfully at all grade levels, but some were not learned at all. Whether a student can understand a given concept is partially a function of the language in which it is presented. Paralleling Piagetian theory, this work raises the question of whether certain concepts can be taught prior to the attainment of a certain level of maturity. However, introducing a concept merely because it is possible to teach it is not necessarily an adequate criterion for determining curriculum content.

Pella and Strauss (1967-1968) reported an extensive study of the



learning of concepts about the biological cell. The authors explored relationships among types of learning as reflected by Bloom's taxonomy, grade level and IQ scores. They found that younger children were not convinced by evidence when the material contradicted the child's logical beliefs about the world. In another statement the authors indicated that lack of basic concepts such as volume and area may contribute substantially to learning difficulties. In another study, Pella and Ziegler (1967-1968) found that mechanical models were effective in teaching selected concepts in grades 2 to 6 with widely varying groups of children.

Trent (1965, p. 229) tested PSSC and traditional physics students on understanding the structure and nature of science; he concluded that "PSSC and traditional physics courses are equally effective in attaining student science understanding as measured by the *Test On Understanding Science*."

*Additional References:* Glass (1965); Hawkins (1966); Ivany (1966); Pella and Voelker (1967-1968); Raven (1967-1968); Smith (1966); Tyler (1968); and Waetjen (1965).

### *Science Instruction for Inquiry*

The new national curriculum projects have emphasized the processes of science. Processes have been discussed under such rubrics as inquiry, enquiry, discovery, and problem solving. Process teaching generates some profound controversies. Shulman's (1968) comparison of the views of Gagné and Bruner indicates Gagné stresses product and Bruner stresses process. Gagné (1933) drew a distinction between discovery and enquiry. He stressed that background knowledge of principles is essential to successfully practicing the techniques of enquiry.

Scott (1966), who is a proponent of inquiry teaching, stated that this teaching procedure encouraged an exploratory attitude in children and resulted in the formulation of questions that were generated within their own cognitive structure. He wrote that inquiry strategies encourage analytical thought and promote intuitive intellectual development. In an examination of the role of inquiry, Rutherford (1964, p. 80) concluded that "the conclusions of science are closely linked with the inquiry which produced them, and, conversely, that the nature of a given inquiry depends upon the topic under investigation. The choice is *neither* facts and laws *nor* inquiry and process; it is *both* facts and laws *and* inquiry and process."

Raun and Butts (1967-1968) attempted to determine the relationship between strategies of inquiry and cognitive and affective behavior. They reported that the use of the strategies appeared to enhance student interests by providing opportunities for assuming an active role and for greater

freedom of action and expression. Evidence indicated that performance in one or more of the selected aspects of inquiry (classifying, observing, using number relations, and recognizing and using time space relations) was positively associated with one or more specific cognitive and affective behaviors: intelligence, divergent thinking, attending, science recall, reading, and attitudinal perception of the potency of science.

Redfield and Atwood (1966) investigated an instructional approach designed to give junior high students an increased capacity to explore physical science problems empirically. Required tasks included: 1) stating a researchable problem, 2) proposing a method of attack, 3) conducting an experiment, 4) using data to make a graph, and 5) interpreting the graph. The students made impressive gains except in stating problems and interpreting graphs. The appropriateness of the criterion for assessing interpretation of the graph is questionable. Students were expected to write the equation of the straight line and to state the upper and lower limits of measurement from which the question was obtained. Can students at this level be expected to have the mathematical sophistication required, and is this an appropriate task for junior high school students?

Tyler (1968) injected a cautionary note about the enthusiasm with which new curricular goals are accepted. He pointed out that the aims of science education have changed several times in the past half century and that although changes in goals are acceptable in the face of new demands and new conditions, it is also possible that new aims are accepted uncritically. Tyler stated (1968, p. 44) that we do not know how helpful the several aspects of science "can be to the layman who is not a professional scientist, nor do we have a carefully considered analysis of the particular attitudes, processes, and products as they relate to the needs of the layman."

Ausubel (1965) is one of the severest critics of the present emphasis on inquiry in science instruction. In an important paper he questioned the philosophical and psychological foundations of most, if not all, of the large-scale curriculum projects of the last decade (1965, p. 259):

Much of this "heuristics of discovery" orientation to the teaching of science is implied by the view that the principal objectives of science instruction are the acquisition of general inquiry skills, appropriate attitudes about science, and training in the operations of discovery. Implicit or explicit in this approach is the belief that the particular choice of subject matter chosen to implement these goals is a matter of indifference (as long as it is suitable for the operations of inquiry), or that somehow, in the course of performing a series of unrelated experiments in depth,

the learner acquires all of the really important subject matter he needs to know.

Ausubel insisted that a science curriculum be a systematic presentation of science as an organized body of knowledge. He stated that the incidental learnings acquired as a by-product of inquiry or discovery pay too little attention to graded and systematically organized content, to substantive and programatic instruction and to practice and feedback variables. Ausubel also refuted the idea that the student is a junior grade scientist; he wrote that deep probing of isolated areas (apart from a systematic presentation of subject matter) merely as a means of enhancing inquiry skills is indefensible.

*Additional References:* George (1965); Gruber (1963); Livermore (1964).

### *Evaluation*

The new curricula have generated many questions of evaluation. Questions about the validity of old measures applied to new curricula are common. Problems of halo effects and in-house evaluations have also raised questions of objectivity and validity. Too often, evaluation has been narrow and lacked historical perspective, psychological and philosophical grounding, and technical proficiency. Welch (1968) made the cogent point that funding of projects by the government, and preparation of materials by scientists, does not insure quality and that evaluation is still indispensable. Welch identified a series of questions which should be asked before new programs are adopted: 1) What is the evidence that the curriculum will achieve the desired objectives? 2) What evidence indicates that the curriculum contributes to the general objectives of science education? 3) What evidence indicates that the program will be successful? He asked additional questions about teacher preparation, student backgrounds, appeal of the course to students, value of the course in future sequences, and teachability.

Many comparative studies of achievement in new and traditional programs have been made. Perhaps the most pervasive impression is the relatively small differences in performance on a wide range of measures. Herron (1966) compared the performance of students in CHEM with those in the traditional chemistry curriculum on the six cognitive levels of Bloom's taxonomy. The results were ambiguous although on the average CHEM students were significantly higher on the "applications" subtest. Neither group made appreciable gains in the two highest cognitive levels—synthesis and evaluation. The low ability group was defined as all students below the 80th percentile on the Iowa Test of Educational Development. The small gains made by all groups as a result of a year's study of chemistry was particularly noteworthy.

Brakken (1965) factor analyzed tests and compared students studying PSSC and traditional physics. Although some small differences were noted, the performances were amazingly similar. This study is particularly interesting because of its approach to comparative evaluation. Its purpose was to determine the degree and manner in which the two physics curricula affected student aptitudes.

Atwood (1967-1968) classified students on the basis of their responses to tests in which the students expressed preferences for memory, fundamental principles, critical questions of information, or practical application. The results showed conclusively that students opting for memory did not perform as well as those who preferred other combinations. It is possible, however, that the less able students expressed a preference for memory. A preference for memory would clearly imply a less demanding performance than other preferences.

*Additional References:* Crumb (1965); Heath and Stickel (1963); Lisonbee (1964); Murphy (1968); Pie and Anderson (1967); Rainey (1964).

### *Toward Unification of Science*

There is a long history of attempts to integrate the recognized fields of science. The need for such synthesis grows more acute. Reporting on the Nuffield Project in England, where there is concern over the large number of children who drop out of science, Keohane (1968, p. 39) espoused a less specialized curriculum which "is . . . essential . . . if the swing from science [student dropouts] which is troubling us so profoundly, is to be reversed. This may seem paradoxical but there is a suspicion, if not firm evidence, that early specialization does, in fact, cause many of the more able pupils to opt for subjects other than science." The Nuffield approach stresses the development of broad unifying themes: interdependence of living things, continuity of life, biology of man, extension of sense perception, movement, harnessing energy, using materials, and the earth and its place in the universe.

Klopfer (1966) built a general case for an integrated science curriculum. He pointed out that the new curricula failed to reflect that the whole of science is an intellectually integrated enterprise. He argued that, although specialized science courses have a place in the secondary school curriculum, for the majority of students science is a fundamental part of general education and should be viewed in the broadest scope possible. He rejected the notion of a science program that perpetuates the artificial separation of science into discrete disciplines and predicted an evolution in science education toward courses that seek to integrate the whole of science.

Selsnick (1963) described work at Ohio State University in develop-



ing a four-year high school course in unified science. A study dealing with a limited aspect of this effort was reported. Selsnick attempted to measure student progress in the traditional and unified programs toward attainment of one major objective, viz., the development of a logical and interrelated picture of the physical world through the mastery of facts, concepts and generalizations of science. Components of a "rational image of the universe" were: phenomena reflecting the interdependencies and interactions of energy, matter and life; forms of matter, natural orderliness revealed in energy and life; things and events perceived in relation to time and space; problems in controlling, predicting and interpreting events transcending single disciplines; and realistic perception of man's relationship with the universe and as a part of the web. Tempered conclusions were reached that unified courses were more successful in helping students obtain the stated objectives.

*Additional References:* Cheldelin and Fiasca (1964); Showalter (1964).

### *Need-Oriented Curricula*

Many workers were concerned with the failure of the new curricula to address a variety of educational needs. Sears (1965) pleaded for conservation education since resource management has facets that touch almost all subjects. He stated that many of the current critical problems in the world cannot be understood unless they are considered in light of the pressures of numbers on a fundamental and finite resource—namely, space. Unfortunately, Sears neglected to note that the new "reform" curricula are virtually devoid of materials directed specifically to such learnings. Brennan (1967) also stressed the need for conservation education. He contended that the preoccupation with atoms and molecules, space and the universe has resulted in neglect of the problems of man in relation to his environment.

A theme in several studies was the neglect of children, especially those of average and below average ability, in the present science curricula. Consideration of the discipline seems to take precedence over the needs of children. Shoresman (1964) suggested that there is a need for some child-focused research that includes such elements as identification of interests and concerns of children at various stages of development, determination of those characteristics of children due to maturation and those due to particular kinds of learning environment, perfection of techniques to ascertain developmental levels of students, more coherent knowledge of readiness, sequence and direction of growth in assimilation of science concepts, and sequencing of activities, materials and subject matter appropriate to different kinds of learners. He recommended that content gradients be derived to serve as guidelines for the development of a sequence of learning experiences.

Watson (1967) contended that the new curricula neglect the average pupil, have limited appeal, and are academically oriented. Particular subject areas have guided curriculum development rather than the interests and abilities of pupils. Each national group has been autonomous and as a result there is no continuity and consistency in the total science curriculum. Bennett (1966) designed a program of large-group, small-group and independent study to effect positive changes in attitudes in low achieving students. He indicated that evidence had been accumulated that students in the program progressed through high school with more confidence in their ability to face life. Thomas (1967) used content from geology and related it to selected historical controversies in a curriculum designed for low-interest students. The use of historical materials resulted in superior achievement for the experimental group.

### *Instructional Resources*

Mertens and Nesbitt (1966) provided an excellent inventory of instructional materials including books, reprints, films, and laboratory exercises. Other information for developing a unit on cytology was included. The resources listed would also have broader usefulness.

Smedley (1963) identified common units of instruction in junior high school science and the apparatus to accompany them. He concluded that material, both content and apparatus, should be selected to provide students with suitable experiences at the beginning of grade 7 and continue throughout the junior high school, and that content and apparatus should be selected to illustrate the broader concepts presented.

*Additional References:* American Institute of Physics (1966); Blackwood (1965); Hutchinson (1964).

### *Summary*

Research during the period of this review shows a desirable tendency toward a broader spectrum of concern, but still lacking are systematic longitudinal studies showing the impact of varied methods and materials on student attitudes, understanding, performance and motivation. Current research seems to be mainly discipline-centered rather than pupil- or learning-centered, and the ends of education appear to be too often subordinated to transitory fashions in educational haberdashery.

### *Bibliography*

- Atkin, J. Myron. Some Evaluation Problems in a Course Content Improvement Project. *Journal of Research in Science Teaching* 1: 129-32; 1963.  
Atkin, J. Myron. Behavioral Objectives in Curriculum Design: A Cautionary Note. *Science Teacher* 35: 27-30; May 1968.

- Atwood, Ronald K. CHEM Study Achievement Among Groups Classified by Cognitive Preference Scores. *Journal of Research in Science Teaching* 5: 154-59; 1967-68.
- Ausubel, David P. An Evaluation of the Conceptual Schemes Approach to Science Curriculum Development. *Journal of Research in Science Teaching* 3: 255-64; 1965.
- Bennet, Leroy C. Livonia's Project Science—A Reality Centered Curriculum. *School Science and Mathematics* 66: 247-54; 1966.
- Brakken, Earl. Intellectual Factors in PSSC and Conventional High School Physics. *Journal of Research in Science Teaching* 3: 19-25; 1965.
- Brennan, Matthew J. Conservation As An Area of Study Appropriate to Science. *Science Teacher* 34: 16-20; Apr. 1967.
- Bronowski, J. Science in the New Humanism. *Science Teacher* 35: 13-16ff; May 1968.
- Bruner, Jerome S. Liberal Education for All Youth. *Science Teacher* 32: 19-21; Nov. 1965.
- Carleton, Robert H. Science Education in the Middle or Junior High School Grades. *Science Teacher* 34: 25-28ff; Dec. 1967.
- Cohen, Robert Sonne. Individuality and Common Purpose: The Philosophy of Science. *Science Teacher* 31: 27-33; May 1964.
- Cronbach, Lee J. Learning Research and Curriculum Development. *Journal of Research in Science Teaching* 2: 204-207; 1964.
- ERIC Information Analysis Center for Science Education. General Bibliography Series #13. Columbus, Ohio: the Center, Dec. 1968a.
- ERIC Information Analysis Center for Science Education. General Bibliography Series #14. Columbus, Ohio: the Center, Dec. 1968b.
- Fox, Fred W. Education and the Spirit of Science—The New Challenge. *Science Teacher* 33: 58-59; Nov. 1966.
- Gagné Robert M. The Learning Requirements for Enquiry. *Journal of Research in Science Teaching* 1: 144-53; 1963.
- Gatewood, Claude. The Science Curriculum Viewed Nationally. *Science Teacher* 35: 18-21; Nov. 1968.
- Gatewood, Claude W. and Obourn, Ellsworth S. Improving Science Education in the United States. *Journal of Research in Science Teaching* 1: 355-99; 1963.
- Gideonse, Hendrik D. Behavioral Objectives: Continuing the Dialogue. *Science Teacher* 36: 51-54; Jan. 1969.
- Herron, J. Dudley. Evaluation and the New Curricula. *Journal of Research in Science Teaching* 4: 159-70; 1966.
- Hungerford, Harold R. Objectives for Student Performance in Junior High School Science: A Viewpoint. *Science Teacher* 35: 35-36; Jan. 1968.
- Karplus, Robert. The Science Curriculum Improvement Study—Report to the Piaget Conference. *Journal of Research in Science Teaching* 2: 236-40; 1964.
- Keohane, Kevin W. Toward an Integrated Teaching of Science. *Science Teacher* 35: 39-43; Oct. 1968.
- Kessen, William. Statement of Purposes and Objectives of Science Education in School. *Journal of Research in Science Teaching* 2: 3-6; 1964.
- Kline, Morris. The Liberal Education Values of Mathematics, Science and Technology for Youth. *Addresses and Proceedings*. Washington, D. C.: National Education Assoc., 1965. Pp. 50-67.
- Klopfer, Leopold. Integrated Science for the Secondary School: Process, Progress and Prospects. *Science Teacher* 33: 27-31; Nov. 1966.
- Lehman, David L. Current Thinking in Adolescent Psychology: Implications for Teaching Science in High School. *Science Teacher* 34: 13-16; Feb. 1967.
- Lockard, J. David. *Sixth Report of the International Clearinghouse on Science and Mathematics Curricular Developments*. College Park: Science Teaching Center, Univ. of Md., 1968.
- Mertens, Thomas R. and Nesbitt, Jerry J. Cytology for In-Service Biology Teachers. *Science Teacher* 33: 16-19; Feb. 1966.
- Montague, Earl J. and Butts, David P. Behavioral Objectives. *Science Teacher* 35: 33-35; Mar. 1968.
- Novak, Joseph. A Model for the Interpretation and Analysis of Concept Formation. *Journal of Research in Science Teaching* 3: 72-83; 1965.
- O'Hearn, George T. *A Review of New Science Curriculum Materials*. Madison, Wis.: Division of Instructional Services, State Department of Public Instruction, 1967.



- O'Hearn, George T. and Pella, Milton O. The Nature and Extent of Societal Implications of Science and Technology in High School Problems of Democracy Textbooks. *Science Education* 51: 212-15; 1967
- Pauling, Linus. The Social Responsibilities of Scientists and Science. *Science Teacher* 33: 14-18; May 1966.
- Pella, Milton O. Science Needed by All. *Science Teacher* 32: 51-52; Sept. 1965.
- Pella, Milton O. and Carey, Russell L. Levels of Maturity and Levels of Understanding for Selected Concepts of the Particle Nature of Matter. *Journal of Research in Science Teaching* 5: 202-15; 1967-68.
- Pella, Milton O. and Strauss, Nyles G. The Relationship Between Concept Attainment and Level of Maturity. *Journal of Research in Science Teaching* 5: 299-310; 1967-68.
- Pella, Milton O. and Ziegler, Robert E. Use of Mechanical Problems in Teaching Theoretical Concepts. *Journal of Research in Science Teaching* 5: 138-50; 1967-68.
- Pella, Milton O.; O'Hearn, George P.; and Gale, Calvin W. Referents to Scientific Literacy. *Journal of Research in Science Teaching* 4: 199-208; 1966.
- Raun, Chester E. and Butts, David P. The Relationship Between the Strategies of Inquiry in Science and Student Cognitive and Affective Behavioral Change. *Journal of Research in Science Teachings* 5: 261-68; 1967-68.
- Redfield, David D. and Atwood, Ronald K. An Investigation of an Experiment Model in the Physical Sciences for the Junior High Student. *Journal of Research in Science Teaching* 4: 217-26; 1966.
- Robinson, James P. Science Teaching and the Nature of Science. *Journal of Research in Science Teaching* 3: 37-50; 1965.
- Rutherford, F. James. The Role of Inquiry in Science Teaching. *Journal of Research in Science Teaching* 2: 80-84; 1964.
- Scott, Norval C. The Strategy of Inquiry and Styles of Categorization. *Journal of Research in Science Teaching* 4: 143-53; 1966.
- Sears, Paul B. Resources Are for Using . . . Wisely. *Science Teacher* 32: 12-17; Sept. 1965.
- Selsnick, Irwin L. The Effectiveness of a Unified Science in High School Curriculum. *Journal of Research in Science Teaching* 1: 302-14; 1963.
- Shoresman, Peter B. Comment and Criticism; Grade Placement—Need For a Shift in Emphasis? *Journal of Research in Science Teaching* 2: 77-79; 1964.
- Shulman, Lee S. Psychological Controversies in the Teaching of Science and Mathematics. *Science Teacher* 35: 34-38, 89-90; Sept. 1968.
- Slawson, Wilber S. Bases of Science Curriculum Development in the Future. *Science Teacher* 35: 22-24; Mar. 1968.
- Smedley, Albert E. Apparatus for Junior High School Science. *Science Teacher* 30: 32-34; Dec. 1963.
- Smith, Herbert A. Educational Research Related to Science Instruction for the Elementary and Junior High School: A Review and Commentary. *Journal of Research in Science Teaching* 1: 199-225; 1963.
- Smith, Herbert A. Appraisal and Projection in Science Education: Some Meditations. *American Biology Teacher* 28: 797-803; 1966.
- Stotler, Donald W. Science, Heuristics and Humanism. *Science Teacher* 32: 28-31; Oct. 1965.
- Summerlin, Lee R. and Craig, Sara P. Evolution of the High School Chemistry Text and Present Implications. *Science Education* 50: 223-33; 1966.
- Sutman, Frank X. Mass Education and the New Science. *Science Education* 50: 494-96; 1966.
- Thomas, Barbara S. Use of a History of Science Case for Low Interest Students. *Science Education* 51: 204-207; 1967.
- Trent, John. The Attainment of the Concept "Understanding Science" Using Contrasting Physics Courses. *Journal of Research in Science Teaching* 3: 224-29; 1965.
- Tyler, Ralph W. Resources, Models and Theory In the Improvement of Research in Science Education. *Journal of Research in Science Teaching* 5: 43-51; 1968.
- Walbesser, Henry H. Curriculum Evaluation by Means of Behavioral Objectives. *Journal of Research in Science Teaching* 1: 296-301; 1963.



- Walbesser, Henry H. Science Curriculum Evaluation: Observations on a Position. *Science Teacher* 33: 34-39; Feb. 1966.
- Watson, Fletcher G. Teaching Science to the Average Pupil. *Science Teacher* 34: 24-26ff; Mar. 1967.
- Welch, Wayne W. Curricular Decisions—How Can Evaluation Assist Science Teachers? *Science Teacher* 35: 22-25; Nov. 1968.

### Additional References

- American Institute of Physics. *Check List of Books and Periodicals for an Undergraduate Physics Library*. New York: the Institute, 1966.
- Blackwood, Paul E. Science Teaching in the Elementary School: A Survey of Practices. *Journal of Research in Science Teaching* 3: 177-97; 1965.
- Cheldelin, Vernon H. and Fiasca, Michael. A Synthesis of the New Curricula in Physics and Chemistry for the Secondary School. *Journal of Research in Science Teaching* 2: 283-87; 1964.
- Crumb, Glen H. Understanding of Science in High School Physics. *Journal of Research in Science Teaching* 3: 246-50; 1965.
- Glass, Bentley. Theory Into Action—A Critique. *Science Teacher* 32: 29-30ff; May 1965.
- George, Kenneth D. The Effect of BSCS and Conventional Biology on Critical Thinking. *Journal of Research in Science Teaching* 3: 293-99; 1965.
- Gruber, Howard E. Science as Doctrine or Thought? A Critical Study of Nine Academic Year Institutes. *Journal of Research in Science Teaching* 1: 124-28; 1963.
- Hawkins, David. Education and the Spirit of Science. *Science Teacher* 33: 18-20; Sept. 1966.
- Heath, Robert W. and Stickel, David W. CHEM and CBA Effects on Achievement in Chemistry. *Science Teacher* 30: 45-46; Sept. 1963.
- Hutchinson, John S. Science Teacher's Adaptable Curriculum. *Journal of Research in Science Teaching* 2: 356-63; 1964.
- Ivany, J. W. George. Psychological Aspects of Structure in Science. *Science Teacher* 33: 36-37; May 1966.
- Lisonbee Lorenzo. The Comparative Effect of BSCS and Traditional Biology on Student Achievement. *School Science and Mathematics* 64: 594-98; 1964.
- Livermore, Arthur H. The Process Approach of the AAAS Commission on Science Education. *Journal of Research in Science Teaching* 2: 271-82; 1964.
- Mahler, Fred. *Appendix and Bibliography to be Used With Life and Earth Science Guides*. Huntsville, Tex.: Cooperative Curriculum Center for Improvement of Educational Opportunities, 1967.
- Murphy, Glenn W. Content Versus Process Centered Biology Laboratories, Part II: The Development of Knowledge, Scientific Attitudes, Problem-Solving Ability, and Interest in Biology. *Science Education* 52: 148-61; 1968.
- National Science Teachers Association. *Bibliography of Science Courses of Study and Textbooks*. Washington, D. C.: the Association, 1968.
- Novak, Joseph. What Should We Teach in Biology? *Journal of Research in Science Teaching* 1: 241-43; 1963.
- Pella, Milton O. and Voelker, Alan M. Teaching the Concepts of Physical and Chemical Change to Elementary School Children. *Journal of Research in Science Teaching* 5: 311-23; 1967-68.
- Pie, Earl L. and Anderson, Kenneth H. Test Achievements of Chemistry Students: A Comparison of Achievement of Students in CHEM, CBA, Conventional and Other Approaches. *Science Teacher* 34: 30-31; Feb. 1967.
- Rainey, Robert G. A Comparison of the CHEM Study Curriculum and the Conventional Approach in Teaching High School Chemistry. *School Science and Mathematics* 64: 539-44; 1964.
- Raven, Ronald J. The Development of the Concept of Momentum in Primary School Children. *Journal of Research in Science Teaching* 5: 216-23; 1967-68.
- Showalter, Victor. Unified Science: An Alternative to Tradition. *Science Teacher* 31: 24-26; Feb. 1964.

- Smith, Herbert A. The Teaching of a Concept: An Illusive Objective. *Science Teacher* 33: 103-12; Mar. 1966.
- Tyler, Ralph W. Analysis of the Strengths and Weaknesses in Current Research in Science Education. *Journal of Research in Science Teaching* 5: 52-63; 1968.
- Waetjen, Walter B. Learning and Motivation: Implications for the Teaching of Science. *Science Teacher* 32: 22-26; May 1965.

## AUTHOR

HERBERT A. SMITH Address: Colorado State University, Fort Collins, Colorado Title: Associate Dean for Education Age: 53 Degrees: B.S., M.A., and Ph.D., Univ. of Nebraska Specialization: Science Education: Administration and Curriculum.



## THE COUNCIL FOR EXCEPTIONAL CHILDREN ANNOUNCES . . . . .

its newest journal, a quarterly publication designed to serve professionals concerned with gifted and handicapped children.

## EXCEPTIONAL CHILD EDUCATION ABSTRACTS . . . . .

is a resource journal containing abstracts stored in the computer file of the CEC's Information Center and features a computer generated Subject Index, Topic Classification Index, Document Classification Index, and Author Index.

The Subject Index is designed to eliminate many hours of library searching by quickly identifying particular documents dealing with a combination of concepts. Researchers can quickly locate research reports dealing with specified variables and practitioners can locate curriculum manuals designed to develop certain skills with a particular type of exceptionality.

Annual subscription rate: first Basic Subscription, \$50; Additional Supplementary Subscriptions, \$25 each. Orders for Supplementary Subscriptions must accompany the Basic Subscription order.

Write: The Council for Exceptional Children, Department 1201 Sixteenth Street, N.W., Washington, D.C. 20036.



### 3: TEACHER EDUCATION IN SCIENCE

Matthew H. Bruce\*

Temple University

In an earlier issue of the *Review*, Burnett (1964) commented on the scarcity of research findings leading toward sound modifications of practice in science teacher education. He gave illustrations of the abundance of "normative studies, particularly those designed to determine college scientific and pedagogic credits earned by science teachers. . . ." To an unfortunate degree this condition still obtains. Furthermore, it appears that a far greater proportion of the literature devoted to science teacher education is concerned with descriptive reporting and what might be termed "reasoned editorializing" than with research findings that might lead to sound modifications of practice.

There are some straws in the wind that indicate an increased movement toward restructuring science teacher education on the basis of valid research evidence, and this is encouraging. More attention is being paid to teacher personality, the effects of teachers' classroom behaviors, and the implications of these effects for science teacher education. Efforts are being made to identify the competencies desirable in a beginning science teacher. There is a growing emphasis on determining more realistic approaches to in-service education and the effects of various components of preparation programs on the eventual classroom performance of science teachers.

A trend is obvious: there is a growing body of research findings responsive to appeals like that made by Watson (1963) for relevant research in areas such as the science teacher's personality and its relationship to teacher education.

#### *Teacher Behaviors, Personality, and Competencies Related to Teacher Preparation*

Although many of the studies exploring teacher behavior and its relationship with other factors do not approach the teachers preparation process directly, they do provide a number of important implications. There seem to be three principal areas among these studies that bear on

---

\*Dr. Chester A. Raun, Temple University, served as consultant to Dr. Bruce on the preparation of this chapter.



teacher education: non-evaluative and evaluative descriptions of behavior or classroom "climate;" the search for relationships between behavior patterns or classroom climates and variables such as student achievement; and the search for relationships between behavior patterns and teacher personality or other personal characteristics.

### *Teachers' Classroom Behaviors*

Parakh (1965) used a modification of the Flanders Interaction Analysis category system to describe verbal patterns in biology teachers. He reported a relatively low percentage of student verbal participation, especially student-initiated contributions, compared with a high percentage of direct verbal teaching procedures employed by most of the teachers in his study. Snider (1965) reported parallel findings in a study of physics teachers. Parakh refrained from making evaluative statements about the teaching behavior patterns observed, but Snider introduced an attempt to relate behavior patterns to teaching effectiveness. This search for a link is also seen in a study by Pankratz (1966). Snider noted a relationship between certain indices of direct teaching behavior and student performance on objective measures of essentially lower-level cognitive learning. Pankratz, however, found that teachers rated as "more effective" on more subjective criteria tended to utilize more indirect teaching activities than did teachers rated as less effective. Although the evidence may be conflicting, the suggestion growing out of such studies for science teacher education is that while preliminary evidence exists that achievement of various types of instructional goals may be related to differential behavior patterns, the preparation of science teachers actually tends toward promoting one type of behavior.

At present the acquisition and analysis of data involved in these studies are expensive, and the consequent small scale of most of the studies now visible places something of a limitation on the generality of the findings. Serious question must be raised about the nature of the criteria of effectiveness in those studies. However, since these studies represent a rather early stage of this type of investigation, one may hope for more sophisticated approaches based on the clues provided by such studies as those noted here.

An investigation of teachers' classroom behavior patterns during the teacher preparation period was reported by McLeod (1967). He noted changes in teaching behavior patterns during the student teaching process, which suggest that the first half of the student teaching period is the time of most rapid change and that patterns of change differ for groups sensitized to examine their own classroom behavior analytically from groups not so sensitized. After comparing his finding with those of Matthews (1966), McLeod concluded that teachers primed to analyze their classroom

behaviors tended to experience more nonrandom changes toward indirect behaviors in their teaching. Although this study suffered from some of the same limitations noted above, it provided a serious implication for teacher education in science. It tended to corroborate Popham's (1965) suggestion that the instructional behaviors of prospective teachers can be modified in planned ways.

Investigations in the area of nonverbal classroom behaviors offered further clues for teacher education. Balzer (1968) provided especially important information suggesting that nonverbal behaviors may contribute a much greater proportion of the total teaching behavior than earlier studies indicated and that nonverbal components which influence the teaching-learning situation occur in a substantial fraction of all teaching behaviors. This suggests a mitigation of the findings of Parakh and others about the very high percentage of direct verbal teaching behavior exhibited by science teachers, and in certain time-based sampling systems (such as the most commonly used Flanders Interaction Analysis and offshoots from it) it suggests some difficulty with interpretations placed on the resultant percentages. Further, Balzer's study raises some doubt about the assumption in some of the studies noted that verbal behavior is representative of total teaching behavior.

Molchen (1967) and McLeod (1967) investigated the student teaching experience and provided evidence of the extent to which student teachers model their behavior patterns after their cooperating teachers in secondary and elementary science teaching.

### *Science Teacher Personality and Other Characteristics*

The goal of an increased number of studies appeared to be the description of science teacher personality or other personal characteristics or the relationship of these to classroom behavior, teaching success and other performance indices.

Craven (1966) compared the critical thinking abilities and understanding of science of prospective science teachers with those of other pre-service and in-service groups. Perhaps his most interesting conclusion is that neither increased critical thinking nor increased understanding of science seems to result from the college study of science. His conclusions should, of course, be qualified by considering the limits of the instruments used for measuring the factors he studied. In particular, the indication regarding "understanding of science" seems suspect because of his use of the Test on Understanding Science, about which there is reasonable doubt as to the appropriateness of its application to persons at the college senior level or persons whose science knowledge is of a fairly high order. To use this instrument which is designed for high school age persons as the basis for a conclusion of "no increase" over the college years is hazardous.

Walberg and Welch (1967) compared personality measures for "innovative" male physics teachers with those for other male high school teachers, including a comparison with other science teachers. Their results suggested that personality differences do exist between physics teachers rated as innovative (as a group) and non-selected teachers, and between these same physics teachers and other science teachers. Although one might reasonably raise a question about almost any definition of "innovative," the strength of the findings in this study and the correspondence with other research findings add weight to the conclusions. The study seems to be a gold mine of clues for further research.

There were comparatively few studies in which the attempt was made to discover relationships between behavioral indices and measures of personality or other personal traits. Evans (1968) suggested, although without much confidence, that some correlations between behaviors and personality traits do exist. He illustrated with greater clarity that the area of teacher personality as related to any behavior or performance index is an exceedingly complex one and needs to be approached with greater sophistication than has typically been applied. The need for increased sophistication in this area, beyond that of first order correlations, has been pointed out by this reviewer before. (Note that the major thrust of the Evans study was toward the development of a reliable category system involving verbal and nonverbal behaviors of biology teachers. This represents an important tool for providing useful information about the teacher education process, particularly in its nonverbal aspect.)

Kimball (1968) developed an instrument to study the understandings of science among scientists and science teachers. He concluded that there was no difference in understanding between these two groups if the science teacher group was restricted to persons whom he defined as qualified (i.e., having completed an undergraduate major in science). Although one might wish to debate certain aspects of the development of the instrument (specifically the use of a model basis offering only a single position and the use as a selection criterion of "... items that ... discriminated in favor of the science graduates. . . ."), one can reasonably see from the description of the development process that Kimball has produced an instrument valid for his purpose, and this warrants confidence in his conclusions. Of special interest regarding the teacher education process is his conclusion about the degree of permanence of the understandings held by the groups examined. He suggested that there is no evidence of significant change as subjects become further removed from the time of their college graduation. His findings do indeed seem to support this. He concluded that the understanding of science must be rather well established by the time students in the groups graduated from college; this conclusion combined with his principal finding of no difference between science and science

teaching groups led him to contend that if greater understanding of science is desired for science teachers, then attention should be given to what happens in undergraduate science instruction. There is merit to this contention. A question which Kimball did not consider directly, but which is touched upon indirectly by his rationale and his findings, concerns the "unqualified" science teachers exempted from his study. His rationale implies that there is a substantial number of these: His findings raise a question about whether whatever level of understanding of science these students have attained at the time of their college graduation can be significantly changed thereafter. This leads to a further question about Kimball's contention of relative permanence of the level of understanding following the undergraduate period: what experience did the subjects of his examination have subsequent to the completion of their baccalaureate education that *might* have brought about change? It appears that the number of subjects involved in his study would not permit approaching this on a "hard data" basis.

### *Desired Competencies*

The question about what the apprentice teacher should be expected to be prepared for remains more or less open. There are three general views of what expectations are reasonable: 1) a practical one (what are teachers actually required to do, or what will administrators accept?); 2) a theoretical one (what do our understandings of the nature of the problem lead us to propose?); and 3) an empirical one (what kinds of teaching behaviors—implying competencies—lead to effective teaching as measured by some standards we can agree upon?). One can hazard the prediction that some synthesis of results from these approaches may eventually lead to acceptable outcomes.

A primary source was utilized in the study by Nelson (1966); he surveyed the actual tasks being performed by a sample of beginning science teachers. The results of this study are somewhat parochial, limited to beginning teachers in one state, but experience suggests that these results are probably typical of what might be found in other geographic areas. Nelson's results are as predictable as they are interesting, but they serve to confirm the subjective observations voiced by others. He reported that teachers were engaging in a wide variety of activities other than just "teaching science"; he also noted discrepancies between what administrators thought were important competencies and what teachers considered were major tasks. Nelson's results emphasize the futility of attempting to produce results meaningful for teacher education by utilizing a single approach. This futility stems not only from the diversity and discrepancies noted by Nelson, but also from the lack in the single approach of a unifying struc-



ture that relates results to the capabilities of teacher education to affect the behaviors of students.

### *Components of Teacher Education Programs*

Perhaps more than any other area of science teacher education, studies of program components suffer from an abundance of "reasoned editorializing" and normative studies. These studies appear as program descriptions, proposals of how things should be done, and surveys of the current state of affairs. For a variety of reasons, such studies occupy space that ought to be devoted to substantive research. (This may be simply a reflection of the scarcity of significant research in this area.) Nearly an entire issue—vol. 3, No. 2, 1965—of the *Journal of Research in Science Teaching* was devoted to descriptions of programs in secondary science teacher preparation at selected institutions. It is nice to know what the other fellow is doing; however, this form of reporting is useful only to the extent that one is willing to compare his qualified, rational judgment to the qualified, rational judgment of someone else and perhaps to act on the comparison.

Status studies are useful in two ways. First, they provide useful feedback when the status is compared with guidelines or other criteria for which there is reasonable expectation that the "status" should reflect a response. Second, they are useful if they document change, or the lack of it, over time.

Anderson (1965) reported a comparison of the status of certain aspects of science teacher preparation to extant guides. He analyzed program components on the basis of nine criterion areas, which were formulated to reflect recommendations for science teacher education then available. Institutions were surveyed within a five-state area. He reported great diversity and numerous deviations from "accepted" guidelines. This is not surprising because at any given time this situation is predictable. Thus, taken by itself, and considering the data sources—questionnaire returns and college program literature, both having uncertain reliability—the study may have little face value, other than local, as a research document. Combined with similar surveys, however, Anderson's study forms a part of a picture over time of the extent to which science teacher preparation programs are changing in response to changing needs and the extent to which they reflect—or fail to reflect—the influence of applicable guidelines.

Reed (1966) focused on a specific component—the science teaching methods course; he presented a study of a popular type. Reed identified commonly utilized practices in methods instruction; he produced a thorough, well-organized catalog of the collective biases of those persons teaching methods courses. Reed's study rises above many similar attempts because he analyzed those conditions in the colleges and universities that impede

the installation of practices which the respondents believed would result in improved programs for teacher education in science. The presentation of these conditions is not unique to Reed's study, but as a part of a longitudinal look at conditions, it adds to the picture of the unwieldiness of the instrument with which the preparation of teachers is attempted.

During recent years an increasing number of attempts were made to examine analytically the actual—as opposed to the idealized—conduct of various components of science teacher education programs. An example of this is the study by Matthews (1966) which was noted earlier. Although it suffered from some of the same limitations of small parochial sampling and problematic assumptions of some of the other behavioral studies, Matthews's study represents an early attempt to analyze objectively a specific component of the science teacher preparation process. In particular, Matthews pinpointed areas in which the student teacher tends to become like his cooperating teacher, as symptomized by verbal behavior patterns.

The introduction of specialized or innovative techniques into the traditional format of teacher education was the subject of some recent research (as well as much description and editorializing.) Goldthwaite (1968) described and evaluated the use of a micro-teaching procedure during student teaching. He searched for relationships between participation in micro-teaching experiences and demonstration performance and rated effectiveness in student teaching. Some relationships were discovered, including a mild surprise in that participants who served as "class members" for the micro-teachers seemed to be rated higher on effectiveness than were the micro-teachers. Interesting as this study is, since it narrows the focus on a specific area and a specific technique, it hinges on an effectiveness index inescapably involving some subjectivity. Furthermore, this study does not in its present state provide much help in predicting the effect the technique may have on the eventual performance of the teachers when they assume the full responsibility of teaching. It does provide clues and suggestions of value in planning future research.

The preparation of science teachers for the junior high school is much talked about but little studied, except for surveys of practice. Perkes (1968) searched for relationships between the background and teaching behaviors of junior high science teachers and the science achievement of their students; his study produced some interesting results. If one accepts his criteria for student achievement—the question raised here is about the goals related to the instruments, not about the instruments themselves—then it would appear that if the achievement goals are essentially factual the number of college credits in science or science teaching methods are not positively related to student achievement. If, however, the goals for student achievement involve more application and interpretation, then the picture revealed by Perkes's findings reverses itself. This is a bit surprising.

Differential results were also obtained from an examination of relationships between teaching behaviors and achievement in the two areas noted. The correlation profiles differed somewhat when broken down by sex and IQ of students, but over all the relation between behaviors exhibited by teachers and achievement by their students seems stronger than that between science course background and achievements; and the differential with respect to the nature of the achievement goals—factual vs. application/interpretation—is a fairly sharp one. Although one may be unhappy about the representation of achievement by standardized test scores, they are what is available, and the study does provide clues for the education of junior high science teachers in the form of differential relationships between two achievement patterns and different patterns of teacher behavior and background in the content of science as well as in methods instruction. The notion of the effect of the recency of the contact with science courses provides an interesting comparison with the assertion by Kimball that the teacher's view of science changes little after his undergraduate education is completed.

### *Elementary Teacher Preparation in Science*

There seems to be less substantive research on the science component of elementary teacher education than of secondary science programs. Descriptions, proposals and normative studies abounded. There were few attempts to discover relationships between components of the teacher education process as it affects science and the eventual classroom performance of elementary teachers in the area of science.

Soy (1964) reported a survey on what happens before or during the preparation of elementary teachers that steers them away from science. Although Soy's study was basically normative and, as is so frequently necessary, localized in its sample of prospective elementary teachers, it nevertheless documents the suspicion of many science educators about effects of elementary science preparation on the elementary teacher. Soy concluded that when prospective elementary teachers arrive at college, their feelings about science are already fairly firmly established. These feelings are not likely to be positive enough to cause these students to elect science courses in college if they can be avoided. Little happens, apparently, in the college program to promote positive feelings. The tie to pre-college science experiences seems strong. If Soy's study were based solely on a questionnaire, one might be tempted to give little weight to her findings, simply because of the high distortion present in much questionnaire data. However, Soy followed up the questionnaire with a substantial number of interviews which lent credibility to her findings. Even so, this study falls short of meeting the need for research findings upon



which valid change can be planned, and in this regard it is similar to many others.

Uhlhorn (1968) described certain pre-student teaching experiences provided for prospective elementary teachers at one institution. He analyzed data resulting from the evaluation of students' presentation of a science lesson. The lesson presentation was part of their instruction in methods of teaching science. Uhlhorn attempted to discover which of several "components" of the presentation influenced that rating given by an instructor, and to measure the influence of ratings from the methods instruction on the final grade in the methods course. He identified some factors associated with success in presentation of a lesson and in terms of the final grade in the course. Uhlhorn noted that this was an "exploratory" study and suggested that further analysis and more information were needed. In this he was quite correct. A study like Uhlhorn's may provide reassurance to an instructor that he is actually doing what he thinks he is doing. But whether it provides useful clues for investigators outside the local situation is questionable because the type of design employed guarantees some form of reassurance, but is incapable of producing any information that was not originally fed into it. When one analyzes the components of a rating, some relationship in terms of those components is bound to emerge. Uhlhorn's study is typical of many others; it is a first order study of use primarily to the author and leaves one hoping the author will continue in pursuit of the larger problem.

The in-service education of elementary teachers with respect to science needs solid research. Reports on the background, attitude and performance of elementary teachers have been numerous for years. Less evident have been studies—as opposed to descriptions—of approaches influencing what happens in elementary classrooms in science. Moser (1964) described a regional approach to improve the science background of elementary teachers. From the practical standpoint, his description of the operational format is useful. Of primary interest, however, is his evaluation of success: an objective test covering the course content was used as a part of the evaluation. Moser's subjects included a large sample (358) drawn from a group of elementary teachers exposed to the course during a two-year period and grouped by teaching level and type of college background. His analysis showed significant gains for all groups. If one assumes that the course was well-designed, the instruction competent and the test relevant, this result is reasonably predictable. More detailed analysis of subgroups provided information about the suitability of the instruction for persons who had varying backgrounds and teaching assignments. Moser attempted another form of analysis using subjective responses of the sample of participants to a series of statements concerning the appropriateness of the material covered, the quality of instruction, etc.



The over-all picture provided by these responses is one of great success in reaching and helping the participants. Of primary concern, however, is the effect that the program had on the classroom activities of the participants. The analysis of the examination has little bearing on this, and the analysis of participants' responses to the evaluative statements touches on classroom activities only obliquely through items related to whether the program materials actually found their way into the participants' classroom or whether the participants' confidence and enthusiasm for teaching science were improved. The value of the report is diminished in the sense that what appears to be a highly workable in-service format for improving both cognitive and affective aspects of the elementary teachers' approach to science is not clearly tied to what happens subsequently in their classrooms.

Rowe and Hurd (1966) reported an approach to evaluation of in-service programs in elementary science. Utilizing a novel approach not easily understood through casual reading, the authors pursued the problem of diagnosing resistance to change in elementary science programs. Their approach was essentially an engineering one and was applied according to the authors' contention that the conditions of evaluation of in-service programs require engineering methods. More specifically, Rowe and Hurd's view is that an in-service program is an attempt to disturb a system in a way anticipated to alter the system in a planned fashion and that evaluation is the comparison of the state of the system afterwards with the desired state. A principal task was the diagnosis of sources of resistance to change in a system. The authors applied the techniques described in the report to a particular in-service program and presented a brief summary of their findings. This is of only passing interest, however. Of primary interest is the analysis of the technique and the demonstration that it did indeed produce the expected results. This report leaves one with the hope that the technique will be applied to other situations and the results reported, for it looks at first reading to be a potent one.

### *Guidelines, Surveys and Summaries*

Several guidelines for the preparation of science teachers have appeared in recent years, but they have tended to be spelled out as program formats in credit hours under generic course names. An exception to this pattern was found in the guidelines for the professional preparation of science teachers produced by a joint effort of the American Association for the Advancement of Science and the Association for the Education of Teachers in Science. The AAAS/AETS (1968) guidelines are notable for several reasons: 1) they represent a high level of cooperative effort between two agencies vitally concerned with the problem; 2) they present a synthesis resulting in a highly adaptable outline of desirable ex-

periences; 3) they are modestly specific without being prescriptive; and 4) they are in harmony with but are not necessarily dependent on earlier content preparation guidelines (such as the 1961 NASDTEC Guidelines). The report specifically presented a number of common elements identifiable in several patterns of teacher preparation, and offered a set of guideline statements adaptable to all the pattern variations noted.

The effect of this type of guideline remains to be seen. According to recent survey reports such as that by Newton and Watson (1968), earlier attempts to stimulate changes in the content and professional preparation of science teachers appear to have had little clearly traceable, substantive results. However, few other attempts at guidelines have involved sufficient breadth of coverage to qualify as much more than local reports or extrapolations of the "reasoned judgment" variety or in many cases even to qualify as research. The AAAS/AETS report does have breadth of coverage and depth as well. It is a synthesis in that it grew from an attempt to build a structure from diversity instead of merely reporting the diversity.

The survey by Newton and Watson led to an essentially non-evaluative report of findings. They surveyed faculty and students on the teacher preparation programs in science at 725 institutions. Topics or program components included program requirements, characteristics of methods courses and student teaching experiences, and reactions from students. In all, the authors surveyed some 15 areas. Although it is impossible in the format of this chapter to review in depth this report, some over-all impressions can be recorded. If one compares the findings of this report with those of earlier attempts to summarize the status of science teacher preparation, such as that by Anderson, one gets the feeling that there has been little general movement in the nature of these programs. The predictable diversity is quite visible in the Newton and Watson (1968) report, not only in the patterns of program arrangement, but in the respondents' outlook regarding what is desirable from the standpoint of the operators of the programs and what is useful from the standpoint of the clientele. Finally, there appear to be large and frequent discrepancies between what is being done and what the respondents believe should be done. These discrepancies are accompanied by a lack of evidence on either the current "doing" or the projected desire to do.

This last impression comes close to summarizing the apparent present state of the preparation of science teachers. However, the evidence is beginning to be collected that should, in the future, lead to those sound modifications of practice which have been so scarce.

## Bibliography

- AAAS/AETS, joint subcommittee from the Cooperative Committee of the American Association for the Advancement of Science and the Association for the Education of Teachers in Science. Guidelines for Content of Preservice Professional Education for Secondary School Science Teachers. *Science Teacher* 35: 85-90; May 1968.
- Anderson, Norman D. *An Analysis of Programs for the Preparation of Secondary School Science Teachers*. Doctor's thesis. Columbus: Ohio State Univ., 1965. Abstract: *Dissertation Abstracts* 26: 1493; No. 3, 1965.
- Balzer, Abe L. *An Exploratory Investigation of Verbal and Non-Verbal Behaviors of BSCS Biology Teachers and Non-BSCS Teachers*. Doctor's thesis. Columbus: Ohio State Univ., 1968. Abstract: *Dissertation Abstracts* 29: 1353A; No. 5, 1968.
- Burnett, R. W. Academic and Professional Preparation of Science Teachers. *Review of Educational Research* 34: 313-21; 1964.
- Craven, Gene F. *Critical Thinking Abilities and Understanding of Science by Science Teacher-Candidates at Oregon State University*. Doctor's thesis. Corvallis: Ore. State Univ., 1966. Abstract: *Dissertation Abstracts*: 27: 125A; No. 1, 1966.
- Evans, Thomas P. *An Exploratory Study of the Verbal and Non-Verbal Behaviors of Biology Teachers and Their Relationship to Selected Personality Traits*. Doctor's thesis. Columbus: Ohio State Univ., 1968. Abstract: *Dissertation Abstracts* 27: 1359A-60A; No. 5, 1968.
- Goldthwaite, D. T. *A Study of Micro-Teaching in the Pre-service Education of Science Teachers*. Doctor's thesis. Columbus: Ohio State Univ., 1968.
- Kimball, Merritt E. Understanding the Nature of Science: A Comparison of Scientists and Science Teachers. *Journal of Research in Science Teaching* 5: 110-20; 1968.
- McLeod, Richard J. *Changes in the Verbal Interaction Patterns of Secondary Science Student Teachers Who Have Had Training in Interaction Analysis and the Relationship of These Changes to the Verbal Interaction of their Cooperating Teachers*. U.S.O.E. Small Contract Project No. 6-8078. Ithaca, N. Y.: Cornell Univ., May 1967.
- Matthews, Charles C. *The Classroom Verbal Behavior of Selected Secondary School Science Student Teachers and their Cooperating Teachers*. Doctor's thesis. Ithaca, N. Y.: Cornell Univ., 1966. Abstract: *Dissertation Abstracts* 28: 144A; No. 1, 1967.
- Molchen, Kenneth J. *A Study of Changes in Intentions, Perceptions and Classroom Verbal Behaviors of Science Interns and Apprentices*. Doctor's thesis. Cambridge, Mass.: Harvard Univ., 1967. Abstract: *Dissertation Abstracts* 28: 2124A-25A; No. 6, 1967.
- Moser, Gene W. An Analysis of the New York State Regional Program in Physical Science for Elementary School Teachers. *Journal of Research in Science Teaching* 2: 85-94; 1964.
- Nelson, T. A. What Work Is Performed By Science Teachers? *Science Education* 50: 346-53; 1966.
- Newton, David E. and Watson, Fletcher G. *The Research On Science Education Survey*. Cambridge, Mass.: Graduate School of Education, Harvard Univ., 1968. (Offset.)
- Pankratz, Roger S. *Verbal Interaction Patterns in the Classrooms of Selected Science Teachers*. Doctor's thesis. Columbus: Ohio State Univ., 1966. Abstract: *Dissertation Abstracts* 27: 1296A; No. 5, 1966.
- Parakh, Jal S. *A Study of Teacher-Pupil Interaction in High School Biology Classes*. Doctor's thesis. Ithaca, N. Y.: Cornell Univ., 1965. Abstract: *Dissertation Abstracts* 26: 6561; No. 11, 1966.
- Perkes, Victor A. Junior High School Science Teacher Preparation, Teaching Behavior and Student Achievement. *Journal of Research in Science Teaching* 5: 121-26; 1968.
- Popham, James. An Experimental Attempt to Modify the Instructional Behavior of Student Teachers. *Journal of Teacher Education* 16: 461-65; 1965.
- Reed, Jack A. *A Description and Evaluation of the Secondary School Science Methods Course in Teacher Preparation Programs in the United States*. Doctor's thesis. Urbana: Univ. of Ill., 1966. Abstract: *Dissertation Abstracts* 27: 692A, No. 3, 1966.
- Rowe, Mary Budd and Hurd, Paul DeH. The Use of Inservice Programs to Diagnose

- Sources of Resistance to Innovation. *Journal of Research in Science Teaching* 4: 3-15; 1966.
- Snider, Ray M. *A Project to Study the Nature of Physics Teaching Using the Flanders Method of Interaction Analysis*. Cooperative Research Project No. S-280. Ithaca, N. Y.: Cornell Univ., Nov. 1965.
- Soy, Eloise M. Attitudes of Prospective Elementary Teachers Toward Science as a Field of Specialty. *School Science and Mathematics* 67: 507-17; 1967.
- Uhlhorn, Kenneth W. Pre-Student Teaching Experiences in Science. *Journal of Research in Science Teaching* 5: 367-72; 1968.
- Walberg, Herbert J. and Welch, Wayne W. Personality Characteristics of Innovative Physics Teachers. *Journal of Creative Behavior* 1: 163-71; 1967.
- Watson, Fletcher G. Academic and Professional Preparation of Science Teachers. *Handbook of Research on Teaching*. (Edited by N. L. Gage.) Chicago: Rand McNally, 1963. Pp. 1031-59.

## AUTHOR

MATTHEW H. BRUCE Address: Temple University, Philadelphia, Pennsylvania Title: Associate Professor of Science Education Age: 38 Degrees: B.S., Indiana State College; M.Ed. and Ph.D., Pennsylvania State Univ. Specialization: Career Determinants and Retention Factors Among Science Teachers; Personality and Other Personal Traits in Science Teachers.



**NEW IN 1969**

# **DEVELOPMENTAL PSYCHOLOGY**

**Editor: Boyd R. McCandless**

*Developmental Psychology* contains articles which represent the broad range of growth and development and their major associated variables. Chronological age as well as sex, socioeconomic status and effects of physical growth variables (including time of maturation and body build) are all considered relevant developmental variables.

Although the adolescent and the aging population appear to deserve particular attention at this time, *Developmental Psychology* indeed covers the entire span. Cross-species articles and articles concerning developmental research with retardates are also included.

The articles in *Developmental Psychology* are not limited to any one approach such as "experiments versus studies" or "cognition versus motivation."

**Bimonthly. One volume per year.**

**Subscription Price: \$10.00 per year.**

**AMERICAN PSYCHOLOGICAL ASSOCIATION**

**1200 Seventeenth Street, N. W.**

**Washington, D.C. 20036**

## 4: CURRICULUM EVALUATION

Wayne W. Welch\*

*University of Minnesota*

Curriculum is frequently defined by educational theorists as the link between society and the schools, the major source of stimuli found in instructional settings, or as pupil behavior pertinent to the goals of the school. According to these definitions curriculum could include everything that happens to a child in school from learning how to accept last Friday's football game loss to the complete sequence of experiences that produced his 800 SAT score. Abramson (1966, p. 388) described the use of broad, general definitions in a *Review* issue on curriculum development and planning: "Clarification of the specific role of curriculum research and evaluation is a continuing problem in the field because of lack of agreement as to the distinctions to be made among curriculum, instructional organization, teaching, and learning."

The absence of consensus necessitates defining the scope of this chapter. *Curriculum* is here defined as a set of materials or planned experiences designed to accomplish certain stated or implied objectives. In science, curriculum has traditionally consisted of syllabi, courses of study, and textbooks. Recently the science curriculum has been dominated by the alphabet programs sponsored by the federal government: BSCS, ESSP, SCIS, etc.

A distinction is made between curriculum and instructional method. This chapter will be limited to studies concerned with the evaluation of curriculum. Discussion of the various teaching strategies is found elsewhere in this issue.

*Evaluation* is the gathering of information for the purpose of making decisions. Curricular decisions generally are made by funding agencies, developers, and eventual users of a program. Evaluation differs from basic research in its orientation to a specific program rather than to variables common to many programs. The objective of educational research is to gain generalizable knowledge about the practice of education; evaluation seeks to provide a basis for making decisions among alternatives (Hemphill, 1969). Thus, evaluation is concerned with questions of utility that involve value and judgment.

\*Drs. J. Myron Atkin, University of Illinois; Hulda Grobman, New York University; and Henry Walbesser, University of Maryland, served as consultants to Dr. Welch in the preparation of this chapter.

Curriculum evaluation serves two important functions: first, it provides a means of obtaining information that can be used to improve a course, and secondly, it provides a basis for decisions about curriculum adoption and effective use. The former is generally called *formative* evaluation; the latter is referred to as *summative* evaluation (Scriven, 1967). The distinction between these two functions is in the manner in which the results are used. Are decisions made about a developing curriculum (formative) or is there need to reach decisions about a curriculum already completed (summative)?

Advances and accomplishments in curriculum evaluation since the last issue of the *Review* devoted to science and mathematics are examined under four general headings: State of the Art, Evaluation in Curricular Development Groups, Evaluation Strategies, and Techniques and Technology. In this chapter, I will describe the general trends in science course evaluation, review recent developments in instruments and methodology, and cite several studies representative of recent efforts to evaluate science curricula.

### *State of the Art*

The five year period covered by this issue of the *Review* was characterized by the existence of several national curricula projects. Groups of scientists, educators, teachers, and others collaborated to rewrite science course materials. These programs were supported mainly by agencies of the federal government; this support initially came from the National Science Foundation (NSF) and more recently it came from the U. S. Office of Education (USOE). A host of new science courses and course materials appeared that espoused goals of "increased scientific literacy," "understanding of the processes of science," and "illustrating the relationship between science and society." In the main, however, these national curriculum projects concentrated on revising course content.

These federally supported curriculum groups dominated the science education scene in recent years; they were the focus of some course evaluation activity. However, they were unwilling or unable to devote the same kind of effort and enthusiasm to evaluation that they gave to development. There was little provision in most of the projects for systematic evaluation either during development (formative) or of the final products (summative). For example, a survey of the *Sixth International Clearinghouse Report of Science and Mathematics Projects* (Lockard, 1968) indicated that only 19 of 68 projects possessed research evidence of success in achieving stated objectives.

It seems, however, that a change is taking place in the evaluation situation. The USOE is supporting curriculum research and evaluation to a greater extent than did the NSF. A "second generation" of curriculum

WELCH

projects has appeared in which more time and effort are being devoted to the problem of evaluation than in earlier projects. An occasional graduate thesis notwithstanding, the period of evaluating a course by using the single criterion of an achievement test peculiar to that course is gone. Because of the influence of the national projects on science education, it will be informative to examine in more detail the evaluation activities of these projects.

### *Evaluation in Curricular Development Groups*

The *International Clearinghouse Report* (Lockard, 1968) described the nature and characteristics of 46 science curriculum projects. The information reported in this document was provided by project directors, and it provides a valuable summary source of data concerning evaluation. This information is summarized in the following series of tables.

The focus of evaluation activity of 46 projects in science is summarized in Table 1. Only two groups carried out all three types of evaluation activities: formative, summative, and basic research. The type of evaluation conducted most by these groups was formative.

TABLE 1

#### NATURE OF EVALUATION ACTIVITY

#### NUMBER OF PROJECTS

Formative	27
Formative and Summative	6
Formative, Summative, and Research	2
None	11
	<hr/>
	46

Table 2 is a breakdown of the various kinds of activity identified in the *International Clearinghouse Report* (Lockard, 1968) as formative evaluation.

TABLE 2

#### TYPE OF FORMATIVE EVALUATION ACTIVITY REPORTED

#### NUMBER OF PROJECTS

Teacher reports (written and verbal)	26
Classroom visitation	15
Student interviews or discussion	9
Questionnaires (structured)	7
Test results	4
Video tapes of classes	3
Professional view of produced materials	2
Computer assisted instruction	1



The total in Table 2 exceeds 35 (the number of projects conducting formative evaluation) because a few projects conducted several different kinds of activity. In general, there was little structure to most feedback activity. Materials were developed, teachers were selected to use them, and they then provided feedback to course developers. During subsequent summers, authors considered the comments made by teachers and students in revising the course. In many projects, this cycle was repeated for several years.

Neither the success of the feedback process in stimulating improvements in curriculum, nor its role in prompting change were extensively studied. Welch (1969) pointed out several practical problems of formative evaluation, e.g., time delays, inability of course authors to incorporate suggested changes, and failure of test items to measure what authors specifically wanted. Although Welch's study was concerned with only one course, it seems likely that the same problems would occur with any other curricular development group. The decision making process in curricular development certainly needs more study than it has received.

Evaluation activities of the curricular development projects can also be characterized by examining the nature of the evaluation instruments that were used. The instruments provide an indication of the type of information evaluators attempted to obtain and identify outcomes these projects were trying to achieve. The instruments used in course evaluation, again summarized from the *International Clearinghouse Report* (Lockard, 1968) are shown in Table 3.

TABLE 3

## EVALUATION INSTRUMENTS USED

## NUMBER OF PROJECTS

Project-developed achievement tests	20
Standardized achievement tests	10
Questionnaires	7
Project-developed cognitive tests	7
<i>Test on Understanding Science</i> (TOUS)	6
Attitude measures	5
Interest measures	3
Biographic information	3
Checklists	2
Critical thinking or reasoning	2
Science process measures	2
Classroom climate inventory	1

The number of different instruments looks impressive in Table 3, but caution is warranted because only eight projects reported conducting

any summative evaluation and only 13 of 46 projects have available evidence of success in achieving stated objectives. The discrepancy between the large number of instruments and the paucity of available results is partially explainable because final reports for several projects have not yet been written. Hopefully, when they are, the contribution of the variety of variables in Table 3 will be realized.

Another way to characterize the evaluation of curriculum projects is to examine basic principles of research design because research and evaluation share many characteristics of method and approach. This is especially true of summative evaluation studies. A tenet of educational research is adequate provision for control group comparisons. Yet only 19 of the 46 science projects discussed earlier reported using control groups. Even among these 19, the method of selection of the control group casts doubt on internal and external validity of results (see Table 4).

TABLE 4

CURRICULUM EVALUATION DESIGN	
NATURE OF CONTROL GROUP USED IN EVALUATION	NUMBER OF PROJECTS
None used	27
Non-randomly selected	11
Randomly selected	4
Method of selection not specified	4
	<hr/> 46

Several of the curriculum development groups were too new to be able to report in 1968 the results of an experimental test of their product. A few projects presented reasons why experimental design was considered inappropriate, particularly during the formative stage of the project. Some questions have been raised about the appropriateness of comparative experimental studies in curriculum evaluation (Cronbach, 1963), but workable alternatives have not yet been presented (Guba, 1969). Hence, evaluators must continue to rely in part on methods of experimentation drawn from basic research.

Only four projects were known to have randomly assigned subjects to either an experimental or control treatment and, thus, can be considered to be true experiments. Campbell and Stanley (1963, p. 172) wrote of the value of true experiments: "[They are] the only means for settling disputes regarding educational practice, the only way of verifying educational improvement. . . ." If this assessment of the importance of

true experimental designs for summative evaluation is correct, then the state of evaluation for curricular development groups is astonishingly inadequate. The information currently available indicates, with a few notable exceptions, that the curriculum reform movement in science, characterized by its enlistment of scholars and scientists, has not been subjected to religious research and evaluation.

### *Evaluation Strategies*

Perhaps the lack of exemplars of successful programs has motivated the considerable evaluation of evaluation that has occurred in recent years. This work has resulted in a number of evaluation paradigms that researchers in science education might find useful. Results of the implementation of these strategies are scarce, but they have provoked considerable interest; several new alternative strategies seem to be evolving.

One evaluation strategy emphasizes measuring the attainment of certain specified objectives. The familiar sequence of defining objectives, planning experiences, and measuring achievement is followed. Perhaps most representative of this approach was the AAAS Commission of Science Education in their elementary science program (Walbesser, 1966). They specified behavioral objectives and developed process measures to determine the proportion of students exhibiting achievement of these objectives.

Atkin (1963) pointed out limitations of this approach, including the claim that evaluation of specified content does not provide for outcomes expected as an outgrowth of many courses. Another difficulty, according to Atkin, is that this approach focuses on rather short-term behavioral changes and tends to obscure the long-term goals. Atkin also warned that stating objectives too early may obscure potential significant outcomes that do not become apparent until later because they are seldom anticipated.

Cronbach (1963, p. 675) quite strongly advocated the use of evaluation for course improvement. He stated, "The greatest service evaluation can perform is to identify aspects of the course where revision is desirable." Flanagan (1969) wrote that formative evaluation represents a much more effective means than summative evaluation for improving education. Stufflebeam (1966) also stressed the formative goal of evaluation as an aid in effective decision making. He particularly emphasized that evaluation should assist with the decisions required by the changes taking place in the developing curriculum.

Another class of evaluation strategies involves decisions that future users of a program are required to make, for example, decisions of adoption, effective use with students of differing needs, and allocation of teaching resources. Scriven (1967) advocated summative evaluation ex-

WELCH

periments to determine the relative effectiveness of new products. He and Stake (1967) elaborated on the summative and formative aspect of evaluation and pointed out the multiplicity of roles and purposes of the evaluator.

In light of the several suggested methods of curriculum evaluation, Tyler and Klein (1967) and Worthen (1968) summarized various approaches and attempted to establish a taxonomy of evaluation designs.

Guba (1969) presented an overview of the evaluation problem and pointed out many of the deficiencies that currently exist in the field of evaluation. He urged a modernization of the theory and practice of the evaluative area. Although it covers more than curriculum evaluation, the sixty-eighth yearbook of the National Society for the Study of Education (1969) provides an excellent discussion of the changes that have been taking place and the many needed changes in educational evaluation.

## *Techniques and Technology*

### *Instrument Development*

Because answers to questions in research and evaluation are eventually provided by data-gathering devices, it is important to have valid and reliable methods. Many national curricular groups gave attention and effort to the problems of achievement testing during the past five years. Nearly half (20 of 46) of the groups surveyed developed, trial-tested, revised, and standardized achievement tests as part of their curricular materials. Typical of the refinement and improvement of the testing instruments were those developed by the Physical Science Study Committee (PSSC), Biological Science Curriculum Study (BSCS), and the Chemical Education Materials Study (CHEM). Several groups developed other general cognitive tests. The Earth Science Curriculum Projects' (1965) Test of Science Knowledge, Part II and the Process of Science Test of the BSCS (1962) were used in their evaluation program to measure understanding of scientific principles and methods. Nearly all of the new achievement tests were written for secondary school students.

Although improvements in achievement testing are worthwhile, they provide an example of the "rich getting richer." Achievement testing procedures are the most advanced of the many outcomes of science instruction that need to be tested. Unfortunately, only a few improvements were noted in the measurement of the more intangible outcomes of science instruction.

An exemplar of the needed innovation in instrument development was the cognitive preference testing technique developed by Heath (1964) for use in evaluating outcomes of the PSSC course. Marks (1967) extended Heath's technique to chemistry and conducted an evaluation of the CBA course using the Cognitive Preference Test: High School Chemistry. Both instruments measured student preference in four areas: (1) memory, (2)



practical application, (3) critical questioning of information, and (4) identification of fundamental principles. The test is innovative, but suffers from several problems. First, neither version appears to be as "discipline free" as one would like. It is doubtful that either CHEM study or Harvard Project Physics would consider the tests appropriate for their use. Second, because the test is ipsative, or "forced-choice," individual scale scores (as opposed to student profiles) are meaningless, and there are statistical problems in performing correlations or factor analyses.

The Test On Understanding Science (TOUS) by Cooley and Klopfer (1961) enjoyed considerable popularity, particularly among doctoral students, and appeared in a few published evaluation studies: Crumb (1965), Trent (1965), and Yager (1968). TOUS explores some important outcomes of science education, but it needs careful revision, re-standardization, and stronger validity evidence. The Welch Science Process Inventory (SPI), Welch and Pella (1968), is designed to measure knowledge of the assumptions, activities, products, and ethics of science. It was validated on the majority opinion of practicing scientists and discriminated among groups of students, teachers, and scientists. The SPI was used as an outcome measure in several curriculum evaluation studies (Welch and Rothman, 1968; Welch and Walberg, 1967-68). Unfortunately, no normative information is available for this inventory.

The semantic differential technique originated by Osgood, Suci and Tannenbaum (1960) was used by several investigators to assess changes in attitudes or interest: Geis (1968), Rothman (1968), and Walberg and Anderson (1968). The discussion of this technique by Geis is an excellent source for researchers considering the semantic differential in curriculum evaluation. Also popular in many studies was the Watson-Glaser Critical Thinking Appraisal (George, 1965; Herron, 1966; Yager, 1968; and Yoesting and Renner, 1969). This test purports to measure one of the general goals of science instruction—critical thinking. There was little success in attaining increases in critical thinking ability that could be attributed to some specified curriculum or teaching method. Yager's study is an exception; he noted significant gains when using a multi-reference approach to physical science versus a single text. General lack of significant results are not too surprising if one examines the Watson-Glaser test carefully. It measures qualities that are learned throughout a lifetime and that seem to be scarcely affected by the short-term instruction that is typically evaluated. Many goals of science education are the cumulative results of several science courses—the cognitive outcomes of only one course may not be detectable with present instruments (Atkin, 1968).

Walberg and Anderson (1967) developed an interesting new instrument for assessing the social climate of learning. Developed as an alterna-

WELCH

tive to classroom observations, the test requires students to state their perceptions of the class on such scales as Intimacy, Formality, Diversity, and Disorganization. Using the Learning Environment Inventory, the investigators found that several different courses affect classroom climate perception differently (Anderson, Walberg and Welch, in press). Also, classroom climate as measured by this instrument was significantly correlated with student cognitive and affective learning (Anderson and Walberg, 1969).

The AAAS program concentrated on the development of non-paper-and-pencil competency measures to determine if students are exhibiting expected behaviors. These competency measures are quite lengthy to administer: the teacher looks for evidence in children's behavior that they have attained the desired process skills (Walbesser and Carter, 1968).

### *Research Design and Data Analysis*

Although multivariate techniques are available, few studies have gone beyond bivariate correlations, *t*-tests, or one-way analyses of variance. Analysis of covariance, multiple and canonical correlation, factor analysis, discriminant analysis, and multivariate analyses of variance and covariance are now within the data analyses capabilities of most universities and offer new possibilities to the curriculum evaluator (Cramer and Bock, 1966).

A few studies used these multivariate techniques and obtained fruitful results. Anderson, Walberg and Welch (in press) used a discriminant analysis with subsequent rotation of axes to identify the differences in courses on the variables of the social climate of learning. They found that the curriculum did have an effect on the students' perceptions of what happened in the classroom.

Finger, Dillon, and Corbin (1965) found in their study of college physics that students who had taken the PSSC course were no more successful than students taking a conventional physics course. Analysis of variance and analysis of covariance were used by these authors. Furthermore, a discriminant analysis was performed that described the characteristics of the groups of the study. The over-all differences in the groups appeared to be mainly attributable to the low performance of students with no prior training in physics. This study can be criticized because a post hoc control group was used, heavy reliance was placed on college grades (generally a very unreliable measure), and it was limited to one university. However, the investigators employed several of the statistical analyses currently available to the curriculum evaluator. Science curriculum evaluation studies that used analyses of covariance to equate for lack of random assignment were few in number (Heath, 1964; George, 1965; Lisonbee and Fullerton, 1964; and Marks, 1967).

The fundamentals of educational research have been known for some

time, but several worthwhile innovations in course evaluation have appeared. Walberg and Welch (1967) successfully used a system of randomized data collection that permits doubling or even quadrupling the number of tests that can be given during a single testing. Randomly-selected halves of each class took different examinations simultaneously. This technique helps to overcome the problem of excessive testing time during curriculum evaluation, yet permits reliable estimates of class means and yields individual data on half of the individuals.

Welch, Walberg, and Ahlgren (1969) showed that it was possible to conduct large-scale course evaluation by enlisting the cooperation of a national random selection of teachers. Their study demonstrated the feasibility of random assignment of teachers to experimental and control groups.

In recent months several booklets appeared that were devoted specifically to the problem of curriculum evaluation in science. Grobman (1968) wrote a review of the approach to evaluation followed by BSCS and other curriculum projects. Lee (1968) reprinted a series of research papers written by staff and graduate students in the Science Education Center at the University of Texas; this series indicates the opportunities in curriculum evaluation and research that a project such as BSCS provides.

### *Results of Some Curriculum Evaluation Studies*

Recognizing the desirability of providing evaluation information in connection with a curricular development project, Atkin and Shoresman (1968) presented a comprehensive summary of the formative and summative evaluation activities conducted by their group. In addition to gathering the informal feedback characteristics of most curricular groups, they examined the effect of the materials on students' perceptions of the nature of science, attitudes toward science, and achievement gains.

The approach to evaluation followed by Harvard Project Physics was described by Welch and Walberg (1968). This group employed a design for evaluation that included formative and summative components, and basic research. More than forty papers were published describing the research and evaluation findings of Harvard Project Physics (Welch, 1969).

Several curriculum studies were conducted by individuals not associated with any of the curriculum groups. However, most of these studies focused on the new curricular materials to determine what effect curriculum alone has on students' achievement, attitude, and behavior. A variety of criterion instruments were used to compare the achievement of students in a "new" course with that of students in a more conventional course. Marks (1967) found that students in the Chemical Bond Approach (CBA) course scored significantly higher on the critical questioning and fundamental principle scale of the Cognitive Preference Test and that conventional chemistry stu-



WELCH

dents scored higher on the memory scale. No significant differences were noted between the two groups on the ACS Cooperative Examination in Chemistry. CBA students scored higher on the CBA final examination. Because neither the method of selection nor the source of the control group teachers was specified (except that they were drawn in the same demographic proportion as the CBA teachers), the results must be viewed with caution. Some teachers chose to teach CBA and others did not; the reasons for the teachers' choices could also explain the findings.

George (1965) examined the contribution of the three BSCS versions (Green, Yellow, and Blue) and conventional biology to increased critical thinking as measured by the Watson-Glaser Critical Thinking Appraisal. The study was carefully conducted and the possibility of problems caused by the limitations of the method of teacher selection was pointed out. Although the study employed an analysis of covariance design, it appears that the number of degrees of freedom should be the number of teachers rather than the number of students since teachers were the sampling unit. The method of teacher selection and the reduction in number of degrees of freedom with the consequent reduction in statistical power raise doubts about the accuracy of the findings.

One way to evaluate a course is to evaluate the textbook. Several methods have appeared for doing this. Newport (1965) used a questionnaire to obtain elementary textbook authors' ratings of nine elementary series. Jacobsen (1965) used an underlining technique to develop readability formulas for chemistry and physics textbooks. He found that readability as measured by this technique was not correlated with popularity of the text for either chemistry or physics.

In a study employing the Watson-Glaser Critical Thinking Appraisal, Herron (1966) found that low and average ability students in conventional courses scored significantly higher on one of the five subtests, Recognition of Assumptions, than did the CHEM study groups. An interesting addition to this study was Herron's attempt to evaluate the two courses on achievement within the different levels of Bloom's taxonomy of objectives in the cognitive domain. He found that CHEM study students scored higher on achievement items determined to measure Application and Analysis, but low ability students in the conventional group scored higher on Analysis than did low ability students in CHEM study.

Two sources of information available to those interested in the results of curriculum evaluation are the newsletters published by the various national projects and the ERIC Center for Science Education at Ohio State University. Most projects use their newsletters for dissemination purposes and usually only summarize results of various evaluation studies. However, the *International Clearinghouse Report* (Lockard, 1968) provides names



and addresses of project directors who can be contacted for additional information.

The ERIC Center has published a number of Science Education Information Reports; one of their general bibliography categories is entitled Evaluation and Educational Objectives (Science Education Information Analyses Center, 1968a, 1968b). These two reports cover a variety of evaluation activities: achievement tests, use of computers in testing, problem solving, etc. Specific references to curriculum evaluation are few in number, however. Only two of ninety listings specify curriculum evaluation as a main topic although inspection of the titles and the list of major ideas suggests that several more could be so listed. For this reason and because of the two-year phase lag that currently exists as the Center tries to catch up on its document processing, I did not find ERIC very useful in preparing this chapter. As the procedures of the Center become more familiar and established, ERIC should become an invaluable aid to curriculum evaluators in science.

### *Summary*

Although the period covered in this review shows signs of increasing emphasis on curriculum evaluation by curricular groups and by individual investigators, there are many areas that seem to need improvement. Several projects have or plan to make evaluation information available to teachers; most have used some sort of feedback system in the development of programs; and a few individual studies have developed new methods and techniques. The literature dealing with curriculum evaluation concentrates mainly on summative studies. The non-comparative but essential formative evaluation seldom appears in the journals, perhaps because it is applied research. Most published studies involve some type of experimental design. However, serious questions can be raised about the lack of statistical rigor and adequate design methodology for these summative studies. Evaluation seems to have advanced beyond the stage of exclusive use of questionnaires (so often criticized in earlier reviews), but there is still inadequate use of the many tools available. Such a fundamental concept as randomization is lacking to the extent that fewer than 5 per cent of the studies can be considered "true" experiments. Problems of multiple *t*-tests, regression toward the mean, inaccurate degrees of freedom for statistical tests, and failure to use covariate adjustment or even simple gains are found throughout the science education literature.

Some strides have been made in instrument development. Many more are needed. Heavy reliance is placed on paper-and-pencil measures; perhaps these can be augmented by new and different instruments. Although teacher-pupil interaction analysis has been used extensively for evaluating instruction, it has been used only sparingly for evaluating curricula (Parakh, 1968;

WELCH

and Snider, 1968). Other techniques that might be used to gather data about the effect of a curriculum include video tapes, pupil dilation or eye movement, film loops, or "unobtrusive measures" as described by Webb et al. (1966).

One conclusion seems obvious. Only at centers where there has been a concentrated effort to investigate many facets of a course or teaching method by a group of researchers does one find any discernible evidence of advancement. Isolated studies by a single graduate student who soon moves on to different interests appear to be making little contribution to the problems of science education. A few centers have sprung up in conjunction with the course development group; these centers have turned out a series of research and evaluation studies that give some depth to the problems. Science educators working with statisticians, educational psychologists, and scientists have made significant advances in several areas. This concentrated group research needs to be encouraged and supported. Isolation in the past has led to fragmentation. Limitations in theory, instrument development, experimental design, data processing, statistical knowledge, and subject-matter competency are difficult for a single investigator to overcome. Several people concerned with common problems and bringing together their own skills and experience appear to offer the best hope for continued improvement of curriculum evaluation in science.

### Bibliography

- Abramson, David A. Curriculum Research and Evaluation. Review of Educational Research 36: 388-95; 1966.
- Anderson, Gary J. and Walberg, Herbert J. Classroom Climate and Group Learning. International Journal of the Educational Sciences 2: 175-80; 1969.
- Anderson, Gary J.; Walberg, Herbert J.; and Welch, Wayne W. Curriculum Effects on the Social Climate of Learning. American Educational Research Journal 6, in press.
- Atkin, J. Myron. Some Evaluation Problems in a Course Content Improvement Project. Journal of Research in Science Teaching 1: 129-32; 1963.
- Atkin, J. Myron. Behavioral Objectives in Curriculum Design. Science Teacher 35: 27-30; May 1968.
- Atkin, J. Myron and Shoresman, Peter B. Elementary School Science Project, Final Report. Urbana: Univ. of Ill., 1968.
- Biological Science Curriculum Study. Process of Science Test. New York: The Psychological Corporation, 1962.
- Campbell, Donald T. and Stanley, Julian C. Experimental and Quasi-Experimental Designs for Research on Teaching. Handbook of Research on Teaching. (Edited by N. L. Gage.) Chicago: Rand McNally, 1963. Pp. 171-246.
- Cooley, William W. and Klopfer, Leo E. Test on Understanding Science. Princeton, N.J.: Educational Testing Service, 1961.
- Cramer, Elliot M. and Bock, R. Darrell. Multivariate Analysis. Review of Educational Research 36: 604-17; 1966.
- Cronbach, Lee J. Course Improvement Through Evaluation. Teachers College Record 64: 672-83; 1963.
- Crumb, Glenn H. Understanding of Science in High School Physics. Journal of Research in Science Teaching 3: 246-50; 1965.
- Earth Science Curriculum Project. Test of Science Knowledge, Part II. New York: The Psychological Corporation, 1965.

- Finger, John A.; Dillon, John A., Jr.; and Corbin, Frederic. Performance in Introductory College Physics and Previous Instruction in Physics. *Journal of Research in Science Teaching* 3: 61-65; 1965.
- Flanagan, John C. The Uses of Educational Evaluation in the Development of Programs, Courses, Instructional Materials and Equipment, Instructional and Learning Procedures, and Administrative Arrangements. *Educational Evaluation: New Roles, New Means*. Sixty-eighth Yearbook, Part II, National Society for the Study of Education. Chicago: Univ. of Chicago Press, 1969. Chapter 10, pp. 221-41.
- Geis, Fred, Jr. *The Semantic Differential Technique as a Means of Evaluating Changes in "Affect."* Doctor's thesis. Cambridge, Mass.: Graduate School of Education, Harvard Univ., 1968.
- George, Kenneth D. The Effect of BSCS and Conventional Biology on Critical Thinking. *Journal of Research in Science Teaching* 3: 293-99; 1965.
- Grobman, Hulda. *Evaluation Activities of Curriculum Projects*. AERA Monograph Series on Curriculum Evaluation, No. 2. Chicago: Rand McNally, 1968.
- Guba, Egon G. *An Overview of the Evaluation Problem*. Paper presented to the American Educational Research Association, Feb. 1969. Bloomington: School of Education, Univ. of Ind. (Mimeo.)
- Heath, Robert. Curriculum, Cognition, and Educational Measurement. *Educational and Psychological Measurement* 24: 239-53; 1964.
- Hemphill, John K. The Relationships Between Research and Evaluation Studies. *Educational Evaluation: New Roles, New Means*. Sixty-eighth Yearbook, Part II, National Society for the Study of Education. Chicago: Univ. of Chicago Press, 1969. Chapter 9, pp. 189-220.
- Herron, J. Dudley. Evaluation and the New Curricula. *Journal of Research in Science Teaching* 4: 159-70; 1966.
- Jacobson, Milton D. Reading Difficulty of Physics and Chemistry Textbooks. *Educational and Psychological Measurement* 25: 449-57; 1965.
- Lee, Addison E. (editor). *Research and Curriculum Development in Science Education: The New Programs in High School Biology*. Univ. of Tex. Publication No. 6720. Austin: The Univ., 1967.
- Lisonbee, Lorenzo and Fullerton, Bill J. The Comparative Effect of BSCS and Traditional Biology on Student Achievement. *School Science and Mathematics* 64: 594-98; 1964.
- Lockard, J. David (editor). *Sixth Report of the International Clearinghouse on Science and Mathematics Curricular Developments*. College Park: Univ. of Md., 1968.
- Marks, Ronald L. CBA High School Chemistry and Concept Formation. *Journal of Chemical Education* 44: 471-74; 1967.
- National Society for the Study of Education. *Educational Evaluation: New Roles, New Means*. Sixty-eighth Yearbook, Part II. Chicago: Univ. of Chicago Press, 1969.
- Newport, John F. An Evaluation of Selected Series of Elementary School Science Textbooks. *Science Education* 49: 479-84; 1965.
- Osgood, Charles E.; Suci, George J.; and Tannenbaum, Percy H. *The Measurement of Meaning*. Urbana: Univ. of Ill. Press, 1957.
- Parakh, Jal S. A Study of Teacher-Pupil Interaction in BSCS Yellow Version Biology Classes. *American Biology Teacher* 30: 841-48, 1968.
- Rothman, Arthur I. The Factor Analysis of a Science-Related Semantic Differential Instrument. *Journal of Educational Measurement* 5: 145-49; 1968.
- Science Education Information Analysis Center. *Science Education Information Report*. Bibliography No. 9. Columbus, Ohio: ERIC Information Analysis Center, May 1968a.
- Science Education Information Analysis Center. *Science Education Information Report*. Bibliography No. 19. Columbus, Ohio: ERIC Information Analysis Center, Dec. 1968b.
- Scriven, Michael. Methodology of Evaluation. *Perspectives of Curriculum Evaluation*. (Edited by Robert E. Stake.) AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally, 1967. Pp. 39-83.
- Snider, Ray M. *A Project to Study the Nature of Physics Teaching Using the Flanders Method of Interaction Analysis*. Doctor's thesis. Ithaca, N.Y.: Cornell Univ., 1966.



WELCH

- Stake, Robert E. The Countenance of Educational Evaluation. *Teachers College Record* 68: 523-40; 1967.
- Stufflebeam, Daniel L. A Depth Study of the Evaluation Requirement. *Theory into Practice* 5: 121-33; 1966.
- Trent, John Henry. The Attainment of the Concept "Understanding Science" Using Contrasting Physics Courses. *Journal of Research in Science Teaching* 3: 224-29; 1965.
- Tyler, Louise L. and Klein, M. Francis. *Recommendations for Curriculum and Instructional Materials*. Paper presented to American Educational Research Association, Feb. 1967. Los Angeles: Univ. of Calif. (Mimeo.)
- Walberg, Herbert J. and Anderson, Gary J. Classroom Climate and Individual Learning. *Journal of Educational Psychology* 59: 414-19; 1968.
- Walberg, Herbert J. and Welch, Wayne W. A New Use of Randomization in Experimental Curriculum Evaluation. *School Review* 75: 369-77; 1967.
- Walbesser, Henry H., Jr. Science Curriculum Evaluation: Observations on a Position. *Science Teacher* 33: 34-39; Feb. 1966.
- Walbesser, Henry H., Jr. and Carter, Heather L. Some Methodological Considerations of Curriculum Evaluation Research. *Educational Leadership* 26: 53-64; 1968.
- Webb, Eugene J. et al. *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago: Rand McNally, 1966.
- Welch, Wayne W. *Harvard Project Physics Research and Evaluation Bibliography*. Cambridge, Mass.: 8 Prescott Street, 1969.
- Welch, Wayne W. *Applied Research in a Curriculum Project: Accomplishments and Limitations*. Paper presented to the American Educational Research Association, Feb. 1969. Minneapolis: College of Education, Univ. of Minn. (Mimeo.)
- Welch, Wayne W. and Pella, Milton O. The Development of an Instrument for Inventorying Knowledge of the Processes of Science. *Journal of Research in Science Teaching* 5: 64-68; 1967-68.
- Welch, Wayne W. and Rothman, Arthur I. The Success of Recruited Students in a New Physics Course. *Science Education* 52: 270-73; 1968.
- Welch, Wayne W. and Walberg, Herbert J. A Design for Curriculum Evaluation. *Science Education* 52: 10-16; 1968.
- Welch, Wayne W. and Walberg, Herbert J. An Evaluation of Summer Institute Programs for Physics Teachers. *Journal of Research in Science Teaching* 5: 105-109; 1967-68.
- Welch, Wayne W.; Walberg, Herbert J.; and Ahlgren, Andrew A. The Selection of a National Random Sample of Teachers for Experimental Curriculum Evaluation. *School Science and Mathematics* 69: 210-16; 1969.
- Worthen, Blaine R. *Toward a Taxonomy of Evaluation Designs*. Paper presented to the American Educational Research Association, Feb. 1968. Columbus: College of Education, Ohio State Univ. (Mimeo.)
- Yager, Robert E. Critical Thinking and Reference Materials in the Physical Science Classroom. *School Science and Mathematics* 68: 743-46; 1968.
- Yoesting, Clarence and Renner, John W. Is Critical Thinking an Outcome of a College Physical Science Course? *School Science and Mathematics* 69: 199-206; 1969.

## AUTHOR

WAYNE W. WELCH Address: University of Minnesota, Minneapolis, Minnesota Title: Associate Director, Educational Research Development Council Age: 34 Degrees: B.S., Wisconsin State Univ.; M.S. (Education), Univ. of Pennsylvania; M.S. (Physics), Purdue Univ.; Ph.D., Univ. of Wisconsin Specialization: Curriculum Evaluation. Measurement, Research in Science Education.



# comparative education review

---

october 1969

*The Social Sciences, Comparative Education, the  
Future, and All That*

W. J. SIFFIN

*Some Heretical Viewpoints on Educational Planning*

C. ARNOLD ANDERSON

*Toward A Theory of Educational Development*

DONALD P. SANDERS

*The Structural Differentiation of Developing  
Educational Systems*

JOSEPH P. FARRELL

*The Impact of a National Service Experience Upon*

*Its Participants: Evidence from Ethiopia* DAVID C. KORTEN & FRANCES F. KORTEN

*Student Radicalism in Japan: A "Cultural Revolution"?*

TOYOMASA FUSE

*Educational Planning and "The System": Myth  
and Reality*

IRA J. WINN

BOOK REVIEWS by BEATRICE BEACH, RANDALL COLLINS, PHILIP H. PHENIX, JOHN  
SMYTH, JOHN A. LASKA, JINDRA KULICH, and WARNER M. MONTGOMERY

PLUS CURRENT BIBLIOGRAPHY OF PERIODICAL ARTICLES AND DOCTORAL DISSERTATIONS  
IN PROGRESS AND COMPLETED

*Forthcoming* A SPECIAL ISSUE devoted to a cross-national study of teachers and  
their classroom activities

The official publication of the Comparative and International Education Society  
Appearing three times annually

**Annual Subscription \$10**

**Student Subscription \$5**

Comparative Education Review, Kent State University, Kent, Ohio 44240

comparative education review

Teachers College Columbia University

New York, New York 10027

## 5: FOREIGN AND INTERNATIONAL PROGRAMS

R. Will Burnett\*

University of Illinois

The last issue of the *Review* devoted to science and mathematics education (1964) did not contain a chapter on international programs. Consequently, the present report includes some selected studies and reports published prior to 1964. Reports of curriculum development are presented in the first section of this chapter. The second section deals with research reports.

### *Curriculum Development*

Numerous reports and materials are available on major curriculum development projects in countries around the globe. Many, perhaps most, of these projects are comparable to the NSF funded curriculum projects in the United States; they are comparable both in the constituents of the working groups (scientists and teachers) and broadly in the materials produced (textual materials, laboratory investigations, materials for teachers, and, sometimes, films and other visual aids). The annual reports of the International Clearinghouse on Science and Mathematics Curricular Developments (Lockard, 1968) are mines of information on these projects. The 1968 report covers over 100 projects from 42 countries and regions, in addition to those from the United States and those sponsored directly by UNESCO. It includes names and addresses of the project director and staff, the nature of the projects, and materials available free or by purchase. The Lockard (1968) report also includes reports on UNESCO pilot projects in chemistry in Asia, physics in Latin America, a world-wide high school science project headquartered in India, and biology in Africa. Bergvall and Nahum (1967) reported on the development of new methods, techniques, and materials for teaching the physics of light in Latin America and on the training of university personnel to teach the program. The eighty-page report includes details about the eight kits of equipment, the development of self-instructional programmed text materials, and the production of 12 short silent films and a half-hour sound film.

\*Drs. Albert V. Baez, Carmel Valley, California; N. E. Bingham, University of Florida; Roger K. Brown, University of Illinois; L. J. Ennever, University of Bristol; Harleen McAda, University of California; and Milton O. Pella, University of Wisconsin, served as consultants to Dr. Burnett on the preparation of this chapter.

Grebecki and Biran (1968) issued a final report on the UNESCO pilot project on new approaches and techniques in biology teaching in Africa. This report provides rather detailed information on the activities of the international working group that carried on the project and on the materials it produced—statements of objectives, teachers' guides, textual materials, and film loops.

The *London Times Educational Supplement* is often the best source of information on English educational developments. Good reports on the Nuffield Science projects (*The Times*, 1967), for example, are found there.

Reports on specific science education centers and their activities are also useful. Unfortunately, these "fugitive" reports are often mimeographed and difficult to obtain. A good example is the report of the Israel Science Teaching Centre (undated) which was established in 1967 to "promote, further and improve the teaching of science in Israel and to change science teaching from the descriptive to the experimental approach."

One of the most active and productive science education centers is located on the campus of the University of the Philippines. The center completed its fourth year in 1968 and is devoted to research and development in science and mathematics education. DeRose (1969) presented a good overview of the center's activities and accomplishments. The newer curriculum materials are being prepared for Philippine students by Philippine scientists and teachers and, with the assistance of American consultants, under the procedural patterns now well-known in the United States. A more extensive and analytical report on this center and its evaluation and assessment program is found in a published newsletter edited by Hernandez (1968a). The reported evaluation program is the equal of the best of those done for the newer curriculum projects in the United States. A pilot test of materials and programs in physics, chemistry, biology, and mathematics included attempts at representative sampling of schools for the experimental tryouts and the matching of some of the experimental classes with control classes. Achievement tests were used in the experimental classes to determine the extent to which the objectives were attained and to diagnose the processes and concepts used in the program. An IQ test for comparability of the groups and a test providing information on the science background of the students were used as pretests.

The testing, to date, indicates that there are no significant differences in over-all composite measures of the teachers of the experimental and control groups, but that the teachers of the experimental groups are, on the average, above the norm in mental ability, science process orientation, and favorable attitudes to personal relationships. The experimental teachers are low on values of conformity. Thirty-five experimental classes were

matched in pairs with as many control classes on the basis of mental ability and subject matter background (school environment was necessarily uncontrolled and, thus, possibly affected the pre- and posttest differences); 2,931 students and 65 teachers were in the matched pairs of classes. Diagnosis of the results disclosed that, in general, the high school students performed much as had been expected. Specifics were reported: e.g., the "low" biology group performed poorly in processes such as interpretation of data and in concepts such as ecological niche. A more detailed report on the biology program of the center, including results of evaluation, was presented by Hernandez (1968b) at a UNESCO conference. This paper focused mainly on the new biology curriculum (adaptation of BSCS) in the Philippines, its impact on students and teachers, and the problems of dissemination and teacher education.

Chaplin (1964) reported on six and one-half years of curriculum development and analysis in Ghana. The work is a major research effort of the University of Ghana and involved 2,000 children, 42 teachers, and three and one-half years of analysis followed by three years of curriculum development work and continued analysis. The earlier work disclosed that older Ghanaian students showed little or no improvement from age nine onward in their ability to analyze simple phenomena or to answer practical phenomenological questions. At the earlier age, nine, the students were "equal in all respects to the American children of New York State" on the basis of identical tests given to the American and Ghanaian children. Even students in Training Colleges showed little superiority to the nine-year-old American children in the tested abilities, although they had a much larger knowledge of facts and definitions than the younger students.

Chaplin felt that the official anxiety to equate Ghanaian syllabi quantitatively with European practice had resulted in frantic but superficial learning. Basing their efforts on the analyses and tentative trials of new materials and approaches, Chaplin and his co-workers developed topics that were teachable entirely by means of the "physical experience of hand and eye in every step of the explanation." Chaplin reported that the experimental topics result in true understanding although the newer method takes from five to eight times as long for understanding to develop as was required for students to memorize under the older and still common Ghanaian methods. Of interest, also, is Chaplin's finding from two years of experimentation that teachers best develop their own understanding by precisely the same sequence of practical experiences that were useful in educating the children.

Russian curriculum development appears to be in as much ferment as that in America—and with comparable emphases. Arsen'ev (1967) reported on the results of the work of a commission established in 1964 by the Presidium of the USSR Academy of Sciences and the RSFSR



Academy of Pedagogical Sciences to raise the scientific level of the content of education in the secondary school of Russia. Results to date include the cutting down of informational material, provision for more thorough scientific explanation of phenomena studied, and the shifting of considerable material from the upper to the lower grades. Leading ideas on the main lines of contemporary scientific development were singled out for each subject (clearly corresponding to conceptual schemes in American parlance). The draft curriculum reported includes many details about biology, chemistry, and physics courses.

*Soviet Education* (1966) contained an interesting and extensive outline of the approved "modern biology" program which is a part of the official 1965 secondary school curriculum revision project in Russia. Molecular biology, genetics, and modern views on evolution (Lysenkoism is apparently completely dead) constitute essential ingredients in the curriculum.

### *Research Reports*

There is a paucity of international research reports available in English. As in American science education, most research in other countries is conducted by candidates for graduate degrees. International and American research suffer from a comparable lack of sophistication, experience, time, and funds available to their researchers. In all countries, most studies are one-shot affairs that are never again pursued by the investigator or by anyone else. Indeed, the infrequency of replication and supporting studies in international research makes action judgments almost impossible. The reports of international and foreign research mainly portray spotty, poorly designed, and truly meager attacks on the fundamental problems that desperately need long-range, sophisticated study.

Els (1959) found a total of 33 studies in science and mathematics education in a screening of more than 1,000 American doctoral dissertations on education in foreign countries. The first of these studies was done in 1905 and the last in 1959. Four were on Canada; nineteen on Europe; four on Africa; and six on Asia. Few articles with or without merit appeared in the research literature during this same period of time.

### *Research in Developing Countries*

During the Depression a considerable amount of American attention in science education was given to the problems of meeting personal and societal needs. As could be expected, developing nations have experienced similarly severe social and personal problems which have been reflected in research efforts. Several studies by nationals of developing nations were designed to direct science education toward personal and societal needs. Hernandez (1960) assumed that science education should contribute

to raising the standard of living and to general social advance in the Philippines; her study was an attempt to determine the direction that science education should take under this assumption. The Hernandez (1960) study was essentially philosophical and intuitive and the conclusions derived inevitably from her basic assumption. It is interesting to compare the purpose and style of this study with her later reports (Hernandez 1968a, 1968b) on the work of the Science Education Center of the University of the Philippines (see previous section in this chapter).

Sangalong (1961) also assumed that science education should reflect personal and social needs. He developed a checklist on rural needs, from a review of the literature on Philippine rural life, and asked the heads of various government agencies involved in rural development to rate the importance of each need on the checklist. From this rating and other sources, Sangalong set up criteria for the development of science learning experiences. The study included four detailed "learning experiences" designed to assist science teachers to help meet the needs (health and increase in farm production, for example) disclosed by the study.

Boulos (1960) and Khalid (1954) carried out studies similar to those of Hernandez and Sangalong. Boulos developed a course of study in biology for Egyptian schools that was based on the "needs of individuals together with the needs of the society. . . ." Khalid analyzed the major social problems of Iraqi society to which science education might make a contribution. He then analyzed the strengths and weaknesses of science education in Iraq and proposed its reconstruction in terms of personal-societal objectives, content, teaching method, examinations, and the training of teachers.

Elashhab (1966) examined the learning process and the nature and structure of science to find the cultural and social problems in the United Arab Republic and to develop a model for science curricula in the secondary schools. His study contained implications about desirable learning experiences and desirable behavioral outcomes. His study also resulted in a set of criteria for stating behavioral objectives, selecting content, organizing learning experiences, and establishing a comprehensive evaluation program.

Hellman (1966) and Phanuel (1966) also conducted studies designed to improve curricula. Hellman examined the social and economic problems of Uganda and determined their implications for instruction in biology. He proposed revised syllabi for the present Cambridge School Certificate and Higher School Certificate Examinations syllabi in biology for Uganda. Phanuel listed 16 characteristics of a good general science curriculum for Madras state, India, and gathered the opinions of secondary school science teachers on the degree to which these characteristics were being satisfied. He also gathered the opinions of science education lecturers

in teacher training colleges. He proposed ways of supplementing the government syllabus and, thus, improve the science offerings in Madras state. Castaño (1967) formulated a set of goals for the sixth grade science program in Manila; she translated these goals into specific content and student behavioral objectives and developed certain instruments and techniques for evaluating pupil growth toward these objectives.

*European, Comparative, and International Studies*

Prince (1968) studied the effect of Western (English) education on children and teachers in New Guinea. His purpose was to determine to what extent the concept of the conservation of physical quantities was the result of "intruding" Western education. A Piagetian science concept test was used. In a typical test item, two identical plasticine balls were displayed. One was then rolled into a cylinder and the testees were asked which of the two objects, if either, had the greater amount of plasticine. The study was carried out with 2,700 children from third grade classes to the third form of high school in three highly diverse districts. Despite the great out-of-school environmental differences, Western-type schooling was a common factor. Comparable results were found by grade but not by age in the three districts; these results led the investigator to conclude that without Western education, the concept of conservation of physical quantities would not be developing "in the way revealed by the survey." This conclusion was supported by the correlation of higher test results with educational experiences with measurement provided in some schools.

Prince's conclusion was further supported by a comparison of two groups of older students, one of which had no background in Western education. Forty-four students at the Awaba Teacher's College had previous formal instruction in English which included length and weight estimation. These students generally did well on the test. But 111 students at the Rintebe Teacher's College, where instruction is in Pidgin and Kate, showed performance "so much below that of the other two groups [the primary grade children and the Awaba students] that one must conclude that conservation of physical quantities is not developed by even a long period of education in Pidgin or Kate" (Prince, 1968, p. 69). In general, the study strongly supported the view that the indigenous Ghanaian culture and even long years of schooling of a non-Western type does not produce concepts of conservation, but that Western education in the same culture, particularly if it deals with quantification, does so.

Poole (1968) used an adaptation of King's Piagetian schedules to measure the influence of environment on scientific concept development. His sample was 150 Hausa tribe students ages 10 and 11 from eight schools in Northern Nigeria. Three schools were urban, three were isolated rural,



and two were in large main road market villages. The samples were randomly selected from the age group in each school. Poole also used a sample of 40 English children with IQ's ranging from 90 to 110; the children were selected randomly from a large junior high school in a southern English city. This sample was used on the assumption that English children would have fair familiarity with the scientific concepts toward which the Hausa groups "are moving"; and to check the assumption that the character and rate of intellectual growth would show cross-cultural variation when measured against the Piagetian stages that imply mastery of "concepts central to Western culture."

The test schedule consisted of 34 items grouped into categories of time and space measurement, prediction of spatial relations, conservation of solids and fluids, grasp of mechanical principles, and attribution of life to inanimate objects. The results were as anticipated for the most part—scientific concept development as tested by the schedules was generally highest among the English children, next highest among the urban Hausa children, and lowest among the isolated rural Hausa children. Analysis of variance showed differences between scores of the Hausa groups and between the Hausa and English samples to be significant in the predicted direction beyond the .01 level. One interesting exception to the trend was the persistence of animism among the English children; e.g., 50 per cent of the English children stated that the sun is alive. This is to be compared with 94 per cent of rural Hausa children and 33 per cent of urban Hausa children who reported the sun to be alive.

Etuk (1967) conducted another Piaget-oriented study of Nigerian children. A standardized interview schedule based on Piaget's tasks was used to investigate the theory that conservation, seriation, and classification develop concurrently. Etuk also examined the relationship of the development of number concepts to intelligence, the effect of sex differences on the development of number concepts, and differences in performances of children from modern and traditional homes. The subjects of Etuk's study were Yoruba-speaking pupils of seven primary schools. Results generally supported Piaget's theory. Etuk found only a slight relationship between intelligence and performance on number tasks. She found slight sex differences and that the children from modern homes performed at a higher level than earlier counterparts from traditional homes.

Raina (1967) used Saxena's 100-item General Science Test in Hindi to study the achievement of 100 elementary school student teachers; he compared the results with those of the Indian standardization group. His sample of student teachers was obtained from the training colleges in two cities in the state of Rajasthan. The Saxena test had been standardized on 2,190 children in the eighth class. Raina's results showed a great difference between the scores of the children used for standardiza-



tion and the teachers in training; the children, on the average, received higher scores. For example, 31 per cent of the children scored from 0-19 (raw scores with the possibilities ranging from 0-100). Fifty-three per cent of the student teachers scored 0-19. Although 48 per cent of the children received scores between 30-89, only 19 per cent of the student teachers achieved in this range. The mean achievement of the children was 26.80 with an S.D. of 12.40; the mean achievement of the student teachers was only 19.80 with an S.D. of 11.00; the difference is highly statistically significant.

Das (1964) compared the problem approach with the conventional lecture-demonstration approach in three high schools in the state of Orissa, India. The problem approach was used in one section and the conventional approach was used in another section of ninth class general science in each of the three schools. The same teacher taught both sections in each school and all teachers were given detailed lesson outlines on the problem approach. Additional training in the problem approach was apparently not provided. A verbal group intelligence test and a pre-nearly the same intelligence and pretest achievement scores were obtained. In these schools, achievement gains of the experimental groups were greater at the .05 level of significance than the conventional groups. In the third school, the experimental group was "much lower" in measured intelligence than the control group. The experimental group was also lower in both the pre- and posttest achievement scores. However, the gain in achievement was about the same (no significant difference) in the experimental and control groups.

Busch (1969) studied the status of science education in the preparatory and secondary schools of Jordan and Syria as indicated by laboratory facilities, the number of full-time vs. part-time teachers, training and experience of teachers, teaching methods (Jordan only), etc. His data were collected by use of a questionnaire distributed by the Ministry of Education in each country and by direct observation of teaching methods in Jordan.

King (1961) developed a Piagetian test of scientific concept development in time and space measurement, spatial relations, mechanical principles, attribution of life including animism, etc. Seventy questions were given to more than 1,800 English school children of ages 5 to 12. On 24 questions there were steady increases in correct responses with age (from about 54 per cent at 6 years to 89 per cent at 11 years). King therefore assumed that these questions could be answered correctly on the basis of general experience and without formal teaching. However, estimations of length, volume, etc. showed very little increase with age,

and the author judged that such questions require formal instruction if appreciable gains are to be made.

Rowlands (1961) used a questionnaire and attitude, interest and vocational choice inventories on a sample of 654 fifteen-year-old English Grammar School boys in a study of differences between prospective scientists, non-scientists, and early school dropouts. The results were analyzed by father's occupation, economic status, number of books in the home, etc. Sociological factors related to type of school, father's occupation, economic status, and cultural level of the home were found to be related to level of aspiration but not with preference for science. Interests, values, and vocational choices were found to be related to both level of aspiration and science preference.

Meyer and Penfold (1961) used subjective interest inventories to measure the leisure science interests, interest in scientific method, and interest in school science topics of 150 students in an East London school. The unweighted total of the three measures was used as the criterion measure of "interest in science." This was then employed as the basis for a rank ordering of presumably related factors such as the father's interest in science. Twenty-eight factors (e.g., sex) were found to be significantly related to the criterion measure and 17 factors (e.g., intelligence) were not significantly related.

Routh (1961) compared chemistry courses in 25 schools selected on the basis of their high reputation for instruction in chemistry in each of the following countries: Canada, Great Britain, New Zealand, and the United States. His study was based on questionnaire responses and on an analysis of the five most widely used textbooks in each country (not necessarily those used in the schools being studied). He ranked the quality of chemistry courses by country as follows: Great Britain, New Zealand, Canada, and the United States.

Atify (1967) used the Science Sequential Tests of Educational Progress (STEP) in grades four, five, and six to determine the achievement of 240 students in three randomly selected schools in Eugene, Oregon, and in three randomly selected schools in Cairo, Egypt. The STEP tests were translated into Arabic for use in Egypt. One must question the validity of the results which showed that the American students were superior to the Egyptian students. Of course they were shown to be. Although the author attributed the results to eight factors (e.g., Egyptian school emphasis on memorization and recitation, pupil-teacher load heavier in Egyptian schools, and facilities inadequate in Egyptian schools), the use of the culturally dependent American STEP test, even in translation, could be enough to account for the differences found. Furthermore, the imposition of American objectives (as inherent in the STEP tests) on

students who have been taught under a different system with different objectives is highly questionable. One wonders what would result if an Egyptian test, based on Egyptian objectives and practices, were translated into English and used for a comparative study. Most comparative studies suffer from the defects of this study. It appears that the best insurance against heavy cultural biases in such investigations would be cooperative studies conducted by both (or all) nationals concerned and abetted by knowledgeable cultural anthropologists.

Harmon (1960) visited 14 schools in Moscow, Leningrad, Kiev, and Irkutsk; he observed instruction and interviewed teachers and administrators. He concluded that Russian science teachers are well trained; they have a light teaching load (18 hours a week; the common method of instruction is lecture-demonstration (only 19 per cent of the schools he visited had student tables adapted for laboratory use); centrally distributed textbooks and strong central control result in much the same work being done in Irkutsk, Siberia as in Moscow; and rote teaching, memory work, and "training" rather than reflective teaching and learning is the common mode.

It is interesting to compare Harmon's status with the views on needed changes expressed by Lerner (1964), a member of the Research Institute of General and Polytechnical Education, RSFSR Academy of Pedagogical Sciences in Russia. Lerner's article first appeared in a Russian journal and certainly reflected an official or at least an officially accepted point of view. Lerner analyzed various teaching approaches and argued for the "research" method to teach students to "extract knowledge" and to learn the elements of research. His lengthy article provided details on this heuristic approach, which he held should be applied broadly and made the object of serious trial and investigation.

Abouseif and Lee (1965) developed "practical" tests in chemistry, biology, and physics. They studied 196 students in English new secondary schools to determine the usefulness of the tests. The requirements for the tests were that they should be mutually independent and valid yet similar for the three subject fields, suitable as a part of the full examination for the Certificate of Secondary Education, and usable by teachers in the field. A panel of science teachers who framed science examinations for the new secondary schools helped in the development of three forms of the tests in each field: single task questions, entire experiments, and a combination of the first two. The tests were based on six hypothetical components of practical ability in science work such as ability to select apparatus and materials for specific purposes, problem solving, and laboratory performance skills. Validity determination of the tests was based on teacher estimates of student skill as observed in actual laboratory performance, on an "attainment test," and on a "scientific ability test."



BURNETT

The tests were piloted in two schools; seven schools in each science subject were used in the main study. Between sixty and seventy students in the fifth form were tested in each field. Comparability of the schools was assessed by the use of a questionnaire. No item analysis, retesting, or reliability measures were reported and presumably were not undertaken. Analysis of the results showed wide differences between schools and between identical test patterns in the three science subjects. This led to the conclusion that identical test patterns probably cannot be used in each field and that the school differences found were probably the result of differences in current stages of development in the Certificate of Secondary Education new type of secondary schools. Such studies are of questionable value. Validity determinations based on other tests of unknown validity and on such informal estimates by several different teachers are hardly worth the effort. The use of a questionnaire to determine comparability of the schools and the lack of hard evidence on comparability, the lack of refinement of items on the test, and the lack of reliability measures all cast doubt on the two major conclusions of the study.

Barrington (1966) compared the achievement of English secondary school physical science students taught by television with that of other students taught by conventional methods. He also used a Likert-type attitude scale to determine student reaction to television instruction. His sample was thirty students in each of five groups from the first and second forms of a four-form entry modern school in Manchester. The standard deviations of general ability of the groups were similar and the initial difference in physical science knowledge between the groups was held constant in the statistical analysis. The B.B.C. series "Discovering Science" was used for five weeks in two television treatments. For one group, the series was integrated with laboratory work. Another group saw the series but did not have laboratory work. These groups were compared with a conventional classroom group following a text program without laboratory work, a group with conventional class and laboratory work, and a group which was taught biology (no physical science) as a control against non-school influences such as home television viewing. All classes were taught by the investigator.

A 100-item achievement test was given to all students prior to instruction, immediately after instruction, and again after a four-month delay. The television non-laboratory group scored higher at the .05 level of significance than did the conventional classroom non-laboratory group. There was no significant difference in achievement between the television-laboratory and the conventionally taught laboratory groups. Each of the two groups taught by television reported favorable attitudes toward television learning; however, the more able the students, the less favorable was their attitude. Barrington noted that the study, while only indicative,



is supported by comparable findings of many studies done in the United States.

### Summation

One emerges from a comprehensive study of international research and development activities with a distinct "I never left home" feeling. Excellent, and sometimes outstanding, research and development activities have been underway. The foregoing listings were winnowed from a larger literature and are both selective and representative. In general, curriculum development is conducted—often sensibly and expertly—without sufficient and continuing assessment, let alone research. Related fundamental problems such as teacher preparation are given comparatively little attention under the pressure to produce learning materials for students. The Ghanaian long-range curriculum development project mentioned earlier and the work of the Philippine science education center are notable exceptions. One would particularly wish that more curriculum efforts were based upon something like the three and a half years of analysis which, in Ghana, essentially preceded and then accompanied the following years of actual curriculum development work. But, then, should not this be a basic principle in curriculum work in the United States as well?

International research suffers from the same ills as those afflicting research in the United States. Long-range investments in research of a comprehensive or even continuing nature are essentially nonexistent. Ground is broken here and there, but fields are never sufficiently worked for a viable crop to be produced. The interesting Piagetian studies reported earlier in this chapter lead one to wish that science education would produce more researchers with Piagetian dedication and skill to work at least a single furrow sufficiently long and well to provide some firmer answers or at least sound hypotheses on which to base science education practices. Despite the large volume of research reports, there is no clear picture of almost anything generalizable. Researchers continue to dig small post holes rather than comprehensibly to work researchable ground.

### Bibliography

- Abouseif, A. A. and Lee, D. M. Evaluation of Certain Science Practical Tests at the Secondary School Level. *British Journal of Educational Psychology* 35: 41-49; 1965.
- Alify, Mohamed Abdel Aziz Eid. *A Comparison of Achievement in Science in the Elementary School in the United States of America and in the United Arab Republic*. Doctor's thesis. Eugene: Univ. of Ore., 1967. Abstract: *Dissertation Abstracts* 29: 140A, 1969.
- Arsen'ev, A. The School and Contemporary Science. *Soviet Education* 9: 3-8; Feb. 1967. Originally published in *Uchitel'skaia gazeta*, Dec. 15, 1966.
- Barrington, H. Teaching of Elementary Science by TV and Other Methods. *British Journal of Educational Psychology* 36: 114-15; 1966.

BURNETT

- Bergvall, Pär and Nahum, Joel. *UNESCO Pilot Project on New Methods and Techniques in Physics Teaching*. Third printing. Paris: UNESCO, 1967.
- Boulos, Sami Ibrahim. *A Course in Biology for the Senior High Schools in Egypt*. Doctor's thesis. Gainesville: Univ. of Fla., 1960.
- Busch, James William. *The Status of Science Education in Preparatory and Secondary Schools in Jordan and Syria, 1963-1965*. Doctor's thesis. Madison: Univ. of Wis., 1969.
- Castañõ, Paulina Maglaque. *Development of Techniques and Procedures for Evaluating the Objectives of the Sixth Grade Science Program in the Elementary Schools in Manila*. Doctor's thesis. Gainesville: Univ. of Fla., 1967. Abstract: *Dissertation Abstracts* 29: 180A, 1969.
- Chaplin, Basil H. G. *The Re-Planning of Junior Science Education in West Africa*. *Science Education* 48: 366-70; 1964.
- Das, Radha Charon. *Problem Development Method: An Experimental Try-Out in the Teaching of Science in Secondary Schools (India)*. *Science Education* 48: 485-90; 1964.
- DeRose, James V. *Philippines Science Education Center Advances with Writing Twenty Courses*. *Science Teacher* 36: 44; Jan. 1969.
- Eels, Walter Crosby. *American Doctoral Dissertations on Scientific and Mathematical Education in Foreign Countries*. *Science Education* 43: 274-75; 1959.
- Elashhab, Gamal A. *A Model for the Development of Science Curricula in the Preparatory and Secondary Schools of the United Arab Republic*. Doctor's thesis. East Lansing: Mich. State Univ., 1966. Abstract: *Dissertation Abstracts* 27: 1540A, 1966-67.
- Etuk, Elizabeth Eme Samson. *The Development of Number Concepts: An Examination of Piaget's Theory with Yoruba-Speaking Nigerian Children*. Doctor's thesis. New York: Columbia Univ., 1967. Abstract: *Dissertation Abstracts* 28: 1295A, 1967.
- Grebecki, Andrzej and Biran, Leonard. *UNESCO Pilot Project on New Approaches and Techniques in Biology Teaching in Africa*. Final Report. Paris: UNESCO, Oct. 30, 1968.
- Harmon, John Millard, Jr. *A Study of Science Teaching in Grades Six through Eight in Selected Ten-Year Schools of the U.S.S.R.* Doctor's thesis. Boston: Boston Univ., 1960.
- Hellman, Robert Alvin. *A Proposed Secondary School Biology Syllabus for Uganda*. Doctor's thesis. New York: Columbia Univ., 1966. Abstract: *Dissertation Abstracts* 27: 2732A, 1966-67.
- Hernandez, Dolores F. *Developing Needs of Science Education in the Republic of the Philippines*. Doctor's thesis. Bloomington: Ind. Univ., 1960.
- Hernandez, Dolores F. *Evaluation Issue*. Newsletter No. 2. Quezon City: Science Education Center, Univ. of the Philippines, Mar. 1968a.
- Hernandez, Dolores F. *The Teaching of Biology in Philippine Schools*. Paper presented at the UNESCO-UNICEF Conference on Science Education Projects in Asia, Bangkok, Nov. 4-18, 1968b. Quezon City: Science Education Center, Univ. of the Philippines. (Typed.)
- Israel Science Teaching Centre. *Report of Organization and Activities*. Jerusalem, Israel: the Centre, Hebrew Univ., undated. (Mimeo.)
- Khalid, Abdulrahman M. *Science Education in the Iraqi Society*. Doctor's thesis. New York: Columbia Univ., 1954.
- King, W. H. *The Development of Children's Scientific Concepts and Interests*. *British Journal of Educational Psychology* 31, Part 1: 1-20; 1961.
- Lerner, I. I. *Acquainting Pupils with Scientific Methods as a Way of Linking Learning with Life*. *Soviet Education* 6: 15-26; 1964.
- Lockard, J. David (editor). *Sixth Report of the International Clearinghouse on Science and Mathematics on Curricular Developments, 1968*. Joint project of the Commission on Science Education, AAAS and the Science Teaching Center, Univ. of Md. College Park: The Univ., 1968.
- Meyer, G. R. and Penfold, D. M. *Factors Associated with Interest in Science*. *British Journal of Educational Psychology* 31, Part 1: 33-37; 1961.
- Phanuel, Mirabai Pushpanathan. *A Plan for Developing an Enriched Curriculum in Science in Madras State, India*. Doctor's thesis. Gainesville: Univ. of Fla., 1966. Abstract: *Dissertation Abstracts* 27: 2952A, 1966-67.

- Poole, H. E. The Effect of Urbanization Upon Scientific Concept Attainment Among Hausa Children of Northern Nigeria. *British Journal of Educational Psychology* 38: 57-63; 1968.
- Prince, J. R. The Effect of Western Education on Science Conceptualization in New Guinea. *British Journal of Educational Psychology* 38, Part 1: 64-74; 1968.
- Raina, T. N. How Well Do Prospective Teachers Know General Science? *Science Education* 51: 234-38, 1967.
- Routh, Charles Joseph. *A Comparison of the Concepts Taught in Secondary School Chemistry Courses in Selected Schools of Canada, Great Britain, New Zealand, and the United States*. Doctor's thesis. Pittsburgh: Univ. of Pittsburgh, 1961.
- Rowlands, R. G. Some Differences Between Prospective Scientists, Non-Scientists, and Early Leavers in a Representative Sample of English Grammar School Boys. *British Journal of Educational Psychology* 31: 21-32, 1961.
- Sangalang, Luz E. *Some Aspects of Science Education for Rural Philippines and Implications for the Education of Teachers for the Community Schools*. Doctor's thesis. Ind. Univ., 1961.
- Soviet Education. On the Draft of a New Program for General Biology. *Soviet Education* 8: 15-27, Jan., 1966. Originally published in *Biologiya v shkole*, No. 4, 1965.
- Subarsky, Zachariah. The Kenya Science Project. *Science Education* 51: 442-46; 1967.
- Times (1967). Nuffield Science Topics. *Times (London) Educational Supplement*: 2711; 15-17-24; May 5, 1967.

## AUTHOR

- R. WILL BURNETT Address: University of Illinois, Urbana, Illinois Title: Professor of Science Education and Chairman of the Department of Secondary and Continuing Education Age: 57 Degrees: B.A., Univ. of Kansas; Ph.D., Columbia Univ. Specialization: Science Education.

## 6: PHILOSOPHICAL AND HISTORICAL BASES OF SCIENCE TEACHING

James T. Robinson\*

*Teachers College, Columbia University\*\**

Previous research reviews in science teaching have not considered philosophical and historical studies. Behavioral studies, curriculum studies, and studies of teaching methods have been most prominent in the literature; but, as Hurd and Rowe (1964) pointed out, such researches have lacked well-developed philosophic starting points and have tended to be contradictory, fragmented, and unpatterned. These teaching and learning studies have appropriated "methods" of science, scientific "principles," science "concepts," etc. as unexamined starting points. However, philosophic studies—studies of methods and principles of science in and of themselves and studies of curriculum design and teaching methods that are consistent with particular methods and principles of science—have been largely neglected.

Several forces have contributed to the science educator's emerging attention to philosophic studies. Scientific knowledge is growing at a rate that makes it impossible for text materials to be completely up-to-date. Growth of knowledge is not only by accretion; rather, it is characterized more appropriately as conceptual reorganization—inert gases are no longer inert, cells are no longer bags of enzymes. A second force relates to science and society. Conceptions of Darwinian evolution, relativity theory, and various other major ideas from the sciences have been and continue to be powerful influences in shaping culture. The tragedy of Oppenheimer and the moral and ethical issues of organ transplant, attest to the inseparability of social problems and scientific knowledge. In the late 1950's science curricula were criticized for being out-of-date and failing to consider science in all of its dimensions. These forces gave impetus to philosophical studies in science education.

### *Problems of Structure in Science*

The idea of "structure of the disciplines" popularized by Bruner, was quickly grasped by those concerned with science curricula. However, it is evident from the studies reported here that structure means different

\*Drs. Walter F. Boldt, University of British Columbia, and F. Michael Connelly, Ontario Institute for Studies in Education, served as consultants to Dr. Robinson on the preparation of this chapter.

\*\*Dr. Robinson is currently with the Biological Sciences Curriculum Study at the University of Colorado in Boulder, Colorado.



things to different investigators. Designers of the curricula sponsored by the National Science Foundation were concerned with structure and with the processes of science. A review of the materials they developed shows that by structure they meant the identification and ordering of concepts and conceptual themes extant in the disciplines. By processes of science they meant primary investigative techniques. Laboratory work was proposed as a major device for teaching these processes (Hurd and Rowe, 1964).

One approach to elucidating structure in science was made by the National Science Teachers Association (1964). Their seven conceptual schemes and five major items in the process of science have both been criticized and defended in the literature (Ausubel, 1965; DeRose, 1965; Glass, 1965; Russell, 1965; Shamos, 1966). The lack of precision, accuracy, and comprehensiveness of the statements, e.g., "every law has its exceptions or its uncertainties," "the process of science has no discernible rules" (NSTA, 1964, p. 19; Novak, 1964, p. 11), contrasts markedly with the substance of the research reviewed here. This lack of precision and comprehensiveness is seductive to many who work in curriculum development because it promises great simplification in the overwhelming task of mastering the manifold realms of scientific knowledge. It creates the illusion of easily grasped solutions rather than hard scholarship for dealing with curriculum problems.

The most penetrating analyses of structure in the literature have been contributed by Joseph J. Schwab. His extensive studies (Schwab, 1951, 1960, 1962, 1963, 1964a, 1964b) are based upon analyses of over 4,000 research reports of scientists and on extensive searches into the origins and assumptions underlying ways of organizing experience. Schwab (1964a) identified three problems of the structure of the disciplines of significance to education: memberships and organizations, syntactical structures, and substantive structures. These problem areas serve as organizing concepts for the discussion below, and Schwab's meanings are considered in the appropriate section.

### *Memberships and Organizations*

Schwab (1964a) argued that different philosophic starting points generate different disciplines and that these disciplines are related to one another differently. Thus, if "science subject matter" and "a simple to complex" order are assumed, the five Comtian disciplines are generated and related to each other in the prerequisite sense: Physics, chemistry, biology, and the social sciences, with mathematics at the bottom. A different philosophic starting point, for example, the "intended ends" of disciplines, generates the Aristotelian disciplines: theoretical, practical, and productive.

The appeal to the Comtian hierarchy is evident in the organization and sequencing of science courses in many schools and colleges. It has not been so evident to those concerned with curricular organization that the Comtian logic is only one of several equally plausible logics. There are several philosophical different ways of determining and organizing disciplines. Each has something it does say and something it does not say about curricula; but, as Schwab indicated, the choice of organization of curriculum cannot be made on the philosophic base alone. Knowledge of the capabilities of students, the ways of learning, and the objectives of instruction must also be considered. (See also Watson, 1963.)

Robinson (1968) investigated six philosophic works by scientist-philosophers of whom three were biologists and three were physicists. By distinguishing explanatory systems that describe from those that predict and with a continuum of correlational and exact procedures, the biological sciences were characterized as having predominately descriptive explanations and physics predominately predictive explanations. Thus, the sciences may be logically ordered in terms of the extensiveness of prediction in their theoretical statements.

### *Syntactical Structure*

Schwab (1964a, 1964b) called the methods of discovery and proof of a discipline its *syntactical structure*. This includes the criteria used for measuring the quality of data, how strictly the canons of evidence are applied, and the pathways by which movement from raw data to conclusions are accomplished. He characterized enquiry in the natural sciences as having short-term syntax (stable enquiry) and long-term syntax (fluid enquiry). *Stable enquiry* proceeds on the basis of accepted substantive principles that guide enquiry in particular fields of science. *Fluid enquiry* produces new conceptions, new models, and special calculus for radically new treatments of problems that can no longer be served by the existing body of scientific knowledge. Fluid enquiry results in periodic revisions of scientific knowledge.

The major concepts of the methodology of the sciences discussed by Robinson (1968) were correlational procedures, exact procedures, rules of correspondence (formal and epistemic), metaphysical principles and constructs. The biological sciences were considered to utilize predominately correlational procedures. Selected sense data are conceptualized into statements that attempt to explain the organization, history, and apparent directiveness of living systems. Many of the concepts within the explanatory statements provide sufficient conditions, but not the necessary conditions demanded by the logical requirements of exact procedures. The polytypic, historical, or functional concepts of biology (e.g., species, mutant, releaser) contrast markedly with the constructs of the physical sciences

(e.g., mass, force, electron) because the latter are linked logically with other constructs and empirically to sense data through rules of correspondence.

Robinson (1968) gave major attention to exact procedures since a strong convergence was identified toward these procedures in the biological and physical sciences. With exact procedures, the movement from sense data to rational thought—to constructs—involves elements of invention and imagination operating within existing theoretical structures. The spontaneity, relative independence, and irreducibility of sense data—tracks on a photographic plate—are brought under methodological control by the establishment of correspondences between sense data and constructs and the subjection of constructs to the demands of metaphysical principles through a circuit of verification. The circuit of verification includes inductive and deductive procedures.

Metaphysical requirements are imposed upon constructs as they are fitted into the theory within which they originated. Constructs must be logically fertile—they must be logically manipulable and obey logical laws. They must be capable of being related to other constructs through formal (logical) or epistemic (empirical) definitions. Constructs that are extensible (Newton's laws included Kepler's laws) have greater utility and are generally more credible than those that are more restricted. Extensibility refers to the range of subject matter included within the construct. Whenever possible, constructs are formulated to generate causal laws. Causality here is considered to be the establishment of an invariable relationship between states, or conditions, of physical systems in which a given state is followed in time by another specific state (e.g., lightning is followed by thunder). The verification of constructs serves two functions: the testing of theory and the testing of the construct.

### *Substantive Structures*

Schwab (1962, 1964b) argued that the principles of enquiry, substantive structures, cannot be discussed in a general way because of their great number and variety. They are designed to fit given subject matters as these subject matters are known at a given time. With the magnitude of sense data exhibited by the subjects of enquiry these principles serve as guides to sort the relevant from the irrelevant. The conceptual nature of substantive principles is evident from the shifts in enquiries that have resulted in changes in scientific knowledge.

Substantive structures are controlled by the opposing criteria of reliability and validity. "Reliability requires that the guiding principles be free of vagueness and ambiguity, that the referents of its terms have unequivocal location and limit, and that measurements or manipulations of

these referents can be made precisely and can be repeated with uniform results" (Schwab, 1964a, p. 27).

Validity requires that the principles guiding enquiry reflect as much as possible the richness and complexity of the subject matter to which they are applied. This criterion is similar to the requirement of extensibility, noted above.

Principles of enquiry may be categorized into one of five forms: reductive principles, holistic principles, rational principles, antiprinciples, and primitive principles (Schwab 1960). Acceptance of specific principles by the scientific community develops as they are tested in the course of enquiry. The relative frequency of principles of enquiry of a particular form may be small in one science and large in another.

*Additional References:* Bronowski (1968); Farre (1966); Neville (1968); Price (1967); Sayvetz (1964); St. John (1968); Vitrogen (1967).

### *Analyses of Structure*

Connelly (1968) based an investigation of research papers in ecology on Schwab's (1960) conception of principles of enquiry. Several hundred research reports were analyzed to identify the problem area and the principles of biology guiding investigations. The higher degree of specificity required to account for the syntactical structure of one particular discipline necessitated the development of a new term. Thus, the term "problem area" was developed for analysis. The "problem area," which eventually became the crucial differentiating term, yielded five classes: classification and taxonomy, energetics, nutrition and metabolism, genealogy, and distribution. Four classes were constructed in terms of principles of biology: antecedent-consequent, structure-function, homeostasis, and regulation. The four classes of principles of biology were identified in ecological research (with two exceptions) in each of the five classes in the problem area. There were, accordingly, 18 significantly different forms of ecological knowledge. The different logical status of the statements was articulated in an enquiry curriculum in terms of problems and principles.

Boldt (1969) utilized Thomas S. Kuhn's (1962) revolutionary view of the growth and development of scientific knowledge as a preliminary stage directed toward the formulation of a theory of classroom teaching and learning. In Kuhn's historical pattern, successive research traditions provide the paradigms through which tacit and explicit knowledge for successful scientific research is acquired, and which are eventually replaced by incommensurable traditions through scientific revolutions. Boldt selected Kuhn's pattern as a source of principles to guide observations of children. He proposed that children's acquisition of views of modern science



might follow an evolutionary pattern of prescientific paradigms, followed by paradigm shifts resulting in new views of phenomena. On a task similar to the Piaget pendulum experiment, Boldt found evidence of paradigm shifts, though many problems of interpretation were encountered.

The tenor of the studies reviewed above is to reject the obsolete philosophies of science that characterize much teaching and learning in the sciences. The scientist as a spectator searching with complete objectivity for absolute truth is replaced by the scientist as a human being in a culture constructing generalizations and testing their consequences by experience. The knowledge produced has a revisionary quality that seems to contradict its apparent permanence and stability, for new formulations replace those currently held and facts lose their significance.

*Additional References:* Anderson (1965); Robinson (1964).

### *Teaching the Sciences*

The preceding studies each contained a recommendation that the characterization of science as a dynamic, human enterprise with a variety of guiding conceptions and procedures be reflected in the materials used for teaching and learning science. In these studies and others, researchers have proposed, produced, and in some cases evaluated materials designed for their recommendations.

The studies I have selected for discussion are grouped into those concerned with teaching structures of science and those concerned with using history of science in science teaching.

#### *Teaching Structures of Science*

Schwab (1962) pointed out the discrepancies between science as he interpreted it and science as reflected in his observations of classroom instruction. He proposed that the current "rhetoric of conclusions"—scientific knowledge conveyed as empirical, literal, and irrevocable truth—be replaced by an enquiring curriculum and classroom. The responsibility of the teacher is, in this context, to impart to the student any art by means of which the student can teach himself. Examples of components of enquiring curricula proposed by Schwab (1951, 1962, 1963) included discussion as a major methodological device, with original scientific papers, idiomatic translation of original papers, and narratives of enquiry as materials. Enquiring laboratory work with three different levels of openness was also proposed. An essential feature for science teaching is to avoid talking about scientific method apart from the content of science. Enquiry, as developed by Schwab, is not presented as a universal logic but as a search that illuminates differences among enquiries within the sciences.

Forty-four "invitations to enquiry" were developed by Schwab (1963). Each was an examination of research reports and included protocol questions for use in discussion. Invitations were grouped into five categories and differed in complexity of data and design. In a widely ignored but important paper, Schwab (1951) provided additional elaboration of the meaning of an enquiring curriculum.

Robinson (1968) proposed an extensive list of understandings that would characterize individuals who are becoming scientifically literate. Procedures for teaching were not a part of his study; however, by advocating that science teaching be consistent with a structure of science, he implied that there are better and poorer materials and methods of science teaching.

Connelly (1968) showed that ecology is characterized by a diversity of patterns of enquiry, and, accordingly, a diversity in its forms of knowledge. He rightly called attention to the lack of specificity provided by curriculum workers in their accounts of scientific methods. He suggested that curricula in ecology should reflect some of the diversity of patterns of enquiry but should not result in a doctrinaire selection of either problem areas or principles of biology. Connelly's proposed curriculum design embodied two ways of teaching science as enquiry: enquiry into enquiry and narrative of enquiry. This proposal was designed to enable the student to recover the meanings embodied in the variety of forms of knowledge by discovering their origins in different problems and different principles. Meaning is enhanced by articulating knowledge with the subject matter it treats, by treating the relations between the facts and outcomes of enquiry, and by treating the effect of the terms of the problem on the pattern of enquiry.

Boldt (1969) proposed that curriculum planning and methods of teaching in elementary science stress the developmental processes of the acquisition of knowledge by children. He suggested that this would be a shift from the current emphasis on displaying a logical structure of the products of scientific research to diagnosing children's paradigms and designing curricula to effect paradigm shifts.

Margenau (1967) criticized contemporary science teaching for its emphasis on induction and its neglect of deduction. In Margenau's view, establishing correspondences between facts and theory leads to the possibility of prediction through deduction. This requires that operational definitions be augmented by formal (logical) definitions that relate construct to construct and permit reasoning. Deductions make possible the prediction of new facts. Emphasis on facts and inductive procedures relegates conceptual illumination to the end of most courses. The press of time generally results in the treatment of laws, theories, and mathematical

equations as facts to be memorized. Emphasis on facts also leads to emphasis on operational definitions. Margenau argued for a balanced emphasis in science teaching on the interplay of pure thought and observation, the formal and the empirical, and the deductive and inductive aspects of science.

Schwab (1962) proposed "narrative of enquiry" as a way to maximize science as enquiry while minimizing teaching as enquiry. The narrative displays conclusions in terms of its evidence and patterns of enquiry in contrast to the traditional rhetoric that merely relates conclusions. He provided an illustration of a narrative of the development of genetics.

Connelly (1967) developed an extensive narrative of enquiry in the role of paleontology in the formulation of theories of hominid evolution. The narrative is organized by Connelly's conception of the nature of paleontological enquiry: methods of analyzing evidence, collecting data, and selecting and interpreting facts. Connelly's narrative reflects a disciplinary structure in three ways. First, the order of evidence reflects the order of discovery of the classes of hominid fossils. Second, a variety of competing accounts of evolutionary patterns is described. Third, different patterns of enquiry are illustrated as different classes of data and different arguments of the same data are described.

*Additional Reference:* Riggsby (1965).

### *Using History of Science in Science Teaching*

Klopfer and Watson (1957) proposed that the use of historical materials could serve as effective means for developing interest in science, appreciation of science, and scientific attitudes. Case histories of the development of major concepts in the sciences were proposed as potential materials for such purposes. In case histories, emphasis is placed on the scientists who were involved, the information available to them, their search for better facts and explanations, and the intellectual and social climate in which they worked. The authors argued that such case studies would be useful in the secondary school and in teacher preparation.

Klopfer and Cooley (1963) investigated the effectiveness of the History of Science Case (HOSC) instructional method in changing student understanding of science and scientists. Two related questions were explored in forty experimental and forty control classes in biology, chemistry and physics: (1) do students who use the HOSC materials achieve greater gains in their understanding of scientists and science than students who do not? and (2) do students who use HOSC materials show as much achievement in the usual science content as students who do not? The Test on Understanding Science (TOUS), Form X, was used as the cri-

terion measure for (1) above, and the Cooperative Biology Test, Form Y; the Cooperative Chemistry Test, Form Z; and the Cooperative Physics Test, Form Z were used as content criterion measures. Students in the experimental classes were taught any two HOSC units selected from six available units by their teachers to best fit their respective programs. The investigators found that the HOSC instructional method was effective in increasing student understanding of science and scientists with little or no concomitant loss of achievement in the usual content of high school science courses.

Thomas (1967) developed a history of science case study, "The Earth's Crust," and used it during 13 days of instruction in physical science classes. She obtained data for 148 students using the case study in three teachers' classes and 175 students in 2 teachers' classes using text book materials. A pretest, posttest design with gain scores on the TOUS test (see Cooley and Klopfer, 1963) was used as the criterion to measure differences in understanding of science and scientists between the two groups. Differences significant at the .05 level for total score and for Scale 1, "The Scientific Enterprise," were found.

Casteel and Yager (1966) reported on the development of the University of Iowa Science and Culture Study. Five units were designed to support a creative nature of science and to demonstrate that scientific knowledge cannot escape the values of the culture in which it is generated. The materials focused on the history of science; the author's guiding philosophy takes issue with the attempt to transfer the operational procedures of the physical sciences to human studies. The major emphasis of the materials is on science within culture, on scientists expressing the values of their culture as they create scientific knowledge and on scientific knowledge affecting culture, penetrating and reshaping it in turn.

Lavach (1967) analyzed several competing scientific theories in an inservice course for 11 teachers in which the historical background of a number of related topics and their subject matter were discussed and reviewed. Based on gain scores on pre- and posttest administration of the TOUS test and in an investigator-designed, twenty-item "Test on the Historical Development of Science," statistically significant differences were found.

Evans (1966) developed a rationale for teaching science through enquiry in a general education course at the college level. He developed text materials in which he sought to follow an historical approach to the philosophy of science reflecting the influence of Kuhn, Toulmin, Duhem, and Hesse. No evaluation—philosophical, historical, or empirical—was conducted.

*Additional References:* Baumel and Berger (1965); Berger (1963); Itzkoff (1966); Kleppinger (1965); Mathias (1962); Winthrop (1965).



### Summary

I examined more than one hundred papers originally identified as being potentially relevant to the topic of this review. Most of these papers were judged to lack sufficient philosophical orientation and analysis to warrant reviewing. Philosophical considerations have a central bearing on the purposes and objectives of science teaching and the materials and methods of instruction. Scholarly contributions to this literature are especially needed in societies in rapid transition.

Scientific knowledge is portrayed as an interlaced constructed system constrained by methodological procedures generated by scientists as they engage in enquiries and develop explanatory systems of the physical world. These studies seek to provide coherent philosophical bases for dealing with structure in science. Their formulations of structures are yet to be challenged, but their influence on teaching practices and curriculum materials is meager.

The differences in patterns of enquiry and in correlational and exact procedures clearly indicate the limitations of statements of conceptual schemes and processes of science at high levels of generality as guides to curriculum development. Conversely, Connelly's successful application of principles of enquiry (substantive structures) to research papers in ecology enhances the potential of Schwab's analytical scheme for future research.

Historical analyses of scientific investigations were proposed as effective means for developing student understandings of the nature of scientific knowledge, the interaction of science and society, and scientists as human beings. Research reported by Klopfer and Cooley (1963) supports this contention if performance on the TOUS test is accepted as evidence.

The critical reader will find the studies of Thomas (1967) and Lavach (1967) so limited in research design [see Campbell and Stanley's (1963) discussion of gain scores and comparison of methods] that they provide little of value for curriculum developers or purchasers.

A second limitation of current studies on the contributions of history of science to science teaching is the lack of sufficiently rigorous comprehension of the underlying philosophical conceptions. In a lucid discussion of the development of the special theory of relativity, Grünbaum (1964, p. 1412) presented a well documented account of his thesis that "philosophy of science is instrumental in illuminating the *genesis* of the conceptual innovations wrought by a particular physical theory." He also showed that textbook accounts of special relativity theory invert "the *logical order* of Einstein's ideas and badly becloud their epistemological anchorage" (Grünbaum 1964, p. 1407).

Schwab (1962) was careful to distinguish materials of instruction (e.g., research papers) from teaching methods (e.g., discussion). A variety

of methods of teaching may be utilized with particular materials. In contrast, Klopfer and Cooley (1963) discussed the History of Science Case *instructional method*. Classifications and distinctions among elements that are parts of teaching will contribute to precision in future research. The materials of instruction may be identical for different groups, but this does not ensure identity of method.

The critical reader may feel that an argument for basing curriculum decisions completely on structures of science has been advocated in the research reported here. Quite the contrary, structures are proposed as necessary, but not sufficient, considerations for curriculum development.

There is no need for additional rhetoric affirming the necessity of teaching process and product in science. Extensive, rigorous philosophical and historical analyses of the sciences from different philosophic starting points are of vital importance to science teaching. This review has examined a few beginnings.

### Bibliography

- Ausubel, David P. An Evaluation of the Conceptual Schemes Approach to Science Curriculum Development. *Journal of Research in Science Teaching* 3: 255-64; 1965.
- Boldt, Walter B. *Application of Thomas S. Kuhn's View of Science to Teaching: An Exploratory Study*. Paper presented to the National Association for Research in Science Teaching, Feb. 1969. Vancouver: Univ. of British Columbia. (Mimeo.)
- Campbell, Donald T. and Stanley, Julian C. Experimental and Quasi-Experimental Designs for Research on Teaching. *Handbook of Research on Teaching* (Edited by N. L. Gage.) Chicago: Rand McNally, 1963. Chapter 5, pp. 171-246.
- Casteel, J. Doyle and Yager, Robert E. Science as Mind-Affected and Mind-Effecting Inquiry. *Journal of Research in Science Teaching* 4: 127-36; 1966.
- Connelly, F. Michael. *Conceptual Structures in Ecology with Special Reference to the Enquiry Curriculum in Ecology*. Doctor's thesis. Chicago: Univ. of Chicago, 1968.
- Connelly, Geraldine. *The Role of Paleontology in the Formulation of Theories of Hominid Evolution—A Teaching Unit*. Master's thesis. Chicago: Univ. of Chicago, 1967.
- Cooley, William W. and Klopfer, Leopold E. The Evaluation of Specific Educational Innovations. *Journal of Research in Science Teaching* 1: 73-80; 1963.
- DeRose, James V. Theory into Action—Contributions to a Unified Science Program. *Science Teacher* 32: 83; May 1965.
- Evans, C. K. *Science as a System of Inquiry—Focus for General Education Programs*. Doctor's thesis. Gainesville: Univ. of Fla., 1965. Abstract: *Dissertation Abstracts* 26: 5783-84; No. 10, 1966.
- Glass, Bentley. Theory into Action—A Critique. *Science Teacher* 32: 29-30, 82-83; May 1965.
- Grünbaum, Adolf. The Bearing of Philosophy on the History of Science. *Science* 143: 1406-12; 1964.
- Hurd, Paul DeHart and Rowe, Mary Budd. Science in the Secondary School. *Review of Educational Research* 34: 286-97; 1964.
- Klopfer, Leopold E. and Cooley, William W. The History of Sciences Cases for High Schools in the Development of Student Understanding of Science and Scientists. *Journal of Research in Science Teaching* 1: 33-47; 1963.
- Klopfer, Leo E. and Watson, Fletcher G. Historical Materials and High School Science Teaching. *Science Teacher* 24: 264-65, 292; Oct. 1957.
- Kuhn, Thomas S. *The Structure of Scientific Revolutions*. Chicago: Univ. of Chicago Press, 1962.

- Lavach, John F. *An In-Service Program in the Historical Development of Selected Physical Science Concepts*. Doctor's thesis. Durham, N.C.: Duke Univ., 1967. Abstract: *Dissertation Abstracts* 28: 4513; No. 11A, 1968.
- Margenau, Henry. *Integrative Education in the Sciences. Main Currents in Modern Thought* 24: 36-41; 1967.
- National Science Teachers Association. *Theory into Action . . . in Science Curriculum Development*. Washington, D.C.: the Association, 1964.
- Novak, Joseph D. Importance of Conceptual Schemes for Science Teaching. *Science Teacher* 31: 10-13; Oct. 1964.
- Robinson, James T. *The Nature of Science and Science Teaching*. Belmont, Calif.: Wadsworth, 1968.
- Russell, James E. Theory into Action—Some Reactions. *Science Teacher* 32: 27-29; May 1965.
- Schwab, Joseph J. Dialectical Means vs. Dogmatic Extremes in Relation to Liberal Education. *Harvard Educational Review* 21: 37-64; 1951.
- Schwab, Joseph J. What Do Scientists Do? *Behavioral Science* 5: 1-27; 1960.
- Schwab, Joseph J. The Teaching of Science as Enquiry. *The Teaching of Science*. (Edited by Joseph J. Schwab and Paul F. Brandwein.) Cambridge Mass.: Harvard Univ. Press, 1962. Pp. 3-103.
- Schwab, Joseph J. (supervisor). *Biology Teachers' Handbook*. New York: John Wiley; 1963.
- Schwab, Joseph J. Structure of the Disciplines: Meanings and Significances. *The Structure of Knowledge and the Curriculum* (Edited by G. W. Ford and Lawrence Pugno.) Chicago: Rand McNally, 1964a. Pp. 6-30.
- Schwab, Joseph J. The Structure of the Natural Sciences. *The Structure of Knowledge and the Curriculum*. (Edited by G. W. Ford and Lawrence Pugno.) Chicago: Rand McNally, 1964b. Pp. 31-49.
- Shamos, Morris H. The Role of Major Conceptual Schemes in Science Education. *Science Teacher* 33: 27-30; Jan. 1966.
- Thomas, Barbara S. Use of a History of Science Case for Low Interest Students. *Science Education* 51: 204-207; 1967.
- Watson, Fletcher. Research on Teaching Science. *Handbook of Research on Teaching*. (Edited by N. L. Gage.) Chicago: Rand McNally, 1963. Chapter 20, pp. 1031-59.

### Additional References

- Anderson, Ronald D. Children's Ability to Formulate Mental Models to Explain Natural Phenomena. *Journal of Research in Science Teaching* 3: 326-32; 1965.
- Baumel, Howard B. and Berger, J. Joel. Teaching from Research Papers, an Approach to Teaching Science as a Process. *Science Teacher* 32: 29-30; Apr. 1965.
- Berger, Melvin. Using History in Teaching Science. *Science Teacher* 30: 24-26; Nov. 1963.
- Bronowski, J. Science in the New Humanism. *Science Teacher* 35: 13-16, 72-73; May 1968.
- Farre, G. L. On the Problem of Scientific Discovery. *Science Teacher* 33: 26-29; Oct. 1966.
- Itzkoff, Seymour W. Physics and History. *Science Education* 50: 485-89; 1966.
- Kleppinger, Edward W. *An Historical and Experimental Approach to the Teaching of Atomic Structure and Sciening*. Doctor's thesis. Toledo, Ohio: Univ. of Toledo, 1965. Abstract: *Dissertation Abstracts* 26: 5789; No. 10, 1966.
- Mathias, Oral B. *An Examination of the Evolution of the First Two Laws of Thermodynamics, Being an Attempt to Discover the Significance of Conceptual Changes Accompanying Their Development*. Doctor's thesis. Kansas City: Univ. of Mo., 1962. Abstract: *Dissertation Abstracts* 28: 544; No. 2A, 1967.
- Nelville, P. No Place for Absolute Truth: The Role of Philosophy and Perception Studies in Science Teaching. *School Science Review* 49: 859-67; 1968.
- Price, Thomas D. *Elementary School Science: A Design for Subject Matter Selection and Organization*. Doctor's thesis. Stanford: Stanford Univ., 1966. Abstract: *Dissertation Abstracts* 27: 3374A-75A; No. 10A, 1967.

- Riggsby, Ernest Duward. *The Scientific Method as Presented in Science Textbooks and as Described by Eminent Scientists*. Doctor's thesis. Nashville: George Peabody College for Teachers, 1964. Abstract: *Dissertation Abstracts* 25: 5144-45; No. 9, 1965.
- Robinson, James T. *An Investigation of Selected Frameworks of Science*. Doctor's thesis. Stanford: Stanford Univ., 1964. Abstract: *Dissertation Abstracts* 25: 1791; No. 3, 1964.
- Sayvetz, Aaron. Subject Matters and Disciplines in the Physical Sciences. *Journal of General Education* 16: 93-102; 1964.
- St. John, Clinton. *Processes and Structure of the Physical Sciences for School Curricula*. Doctor's thesis. Toronto: Univ. of Toronto, 1968.
- Vitrogen, David. Origins of the Criteria of a Generalized Attitude Toward Science. *Science Education* 51: 175-86; 1967.
- Winthrop, Henry. Teaching the History of Science: Some Suggested Innovations in Method. *Science Education* 49: 420-32; 1965.

## AUTHOR

JAMES T. ROBINSON Address: Teachers College, Columbia University, New York, New York Title: Associate Professor of Natural Sciences Age: 46 Degrees: A.B. and M.S., Whittier College; Ed.D., Stanford Univ. Specialization: Philosophy of Science and Science Teaching; Testing in High School Biology.



# Books from Blaisdell

**John I. Goodlad, University of California, Los Angeles**

**School, Curriculum, and the Individual**

Hardcover. \$8.50

**Bruce R. Joyce, Teachers College, Columbia University**

**Alternative Models of Elementary Education**

Hardcover. \$5.75

**Raymond G. Kuhlen**

**Studies in Educational Psychology**

Paperbound. \$6.50

**Lois N. Nelson, San Francisco State College**

**The Nature of Teaching**

Paperbound. \$6.50

**David A. Payne, Syracuse University**

**The Specification and Measurement of Learning Outcomes**

Paperbound. \$2.95

**David A. Payne, Syracuse University, and Robert F. McMorris  
State University of New York at Albany**

**Educational and Psychological Measurement**

Paperbound. \$6.50

**David B. Tyack, University of Illinois, Urbana**

**Turning Points in American Educational History**

Paperbound. \$6.75



**Blaisdell Publishing Company**

A Division of Ginn and Company

**275 Wyman Street**

**Waltham, Massachusetts 02154**

## 7: CURRENT RESEARCH IN MATHEMATICS EDUCATION

Thomas A. Romberg\*

*University of Wisconsin*

Current research in mathematics education can be characterized as large in quantity, poor but improving in quality, and diverse. During the past five years over a thousand studies of mathematics instruction have been reported. The poor quality of most studies can be attributed to too much interest in mathematical components and too little concern for experimental design, measurement, or analysis. The diversity of this immense number of studies was a major problem in preparing this review.

The school level categorization of past *Review* issues was not followed. The intent of the present organization is to highlight certain problem areas that cut across school levels. (For school level reviews see Glennon and Callahan, 1968, and Willoughby, 1969.) For this issue research reports have been categorized into the following eight areas: 1) mathematical learning from an association learning framework; 2) mathematical learning from an activity learning framework; 3) mathematical problem solving and creative behavior; 4) mathematics teaching; 5) the effectiveness of instructional programs; 6) the association of learner characteristics with mathematical achievement; 7) attitudes toward mathematics; and 8) the evaluation and measurement of mathematics achievement.

These eight areas reflect the stimulation of mathematics education by the new courses developed in the early 1960's and the growing interest and involvement of mathematics educators and psychologists in systematic studies of the learning and teaching of mathematics. At the beginning of this decade, the United States was swept with enthusiasm for improving scientific and mathematical training programs. Now as the decade draws to a close, salesmanship has given way to questioning and, in some cases, to careful inquiry. This does not imply that the objectives of the reform movement were failures; there is no question that the content of mathematics courses needed to be updated. However, the educational process is complicated. Changes in content alone were not sufficient to produce drastic changes in mathematical learning. There is considerable agreement about what should be taught, but not about how it should be taught

---

\*Mrs. Deborah M. Stewart, University of Wisconsin, served as consultant to Dr. Romberg on the preparation of this chapter and as a reviewer for the other chapters.

(Begle, 1968). The implications of the content changes for children learning in various settings was not and is not clearly understood (see Heath, 1964).

The challenge of the early curriculum reforms to the orthodox views of curricula and learning was significant. The most astute educational scholars quickly saw that the advocated changes were more than an exchange of new concepts for old. "The greatest novelty in the new curricula," as Cronbach (1965, p. 121) pointed out, "is not the content, not the instructional methods, not the grade placement of topics [but] . . . the objective from which all else stems." The essential curricular aim was to have mathematics taught as a discipline—a system of thought as the specialist knows it, which includes the systematic contents of the discipline and its procedures.

The potentially most important avenues of research led to attempts to examine theoretically the learning of mathematics as a discipline. This research, however, did not follow a simple pattern. Researchers followed such leads as the structure of mathematics so convincingly portrayed by Bruner (1960), the hierarchical nature of mathematics championed by Glaser (1968), or the discovery process of learning central to many new curricula and recently criticized carefully by a number of scholars (see Shulman and Keislar, 1966). The categorization scheme chosen for this *Review* is an attempt to organize the recent activities into reasonable areas for summarization. The two categories on mathematical learning were suggested by Bourne (1966). He characterized contemporary trends in psychological learning theories either as associational—those picturing students as recipients of information—or as hypothesis testing—those picturing students as active participants in the process of learning. Hiemer and Kieren summarized the important research in these areas in their chapters.

Kilpatrick summarized research on mathematical problem solving and creativity in his chapter. To many nonmathematicians, mathematics is simply a set of concepts and skills to be acquired or mastered. However, to mathematicians, the ideas of the discipline have always been a framework from which they could create new ideas or new solutions to problems. The last chapter by James Fey is a summary of research on the teaching of mathematics. Due to space limitations these chapters include only the most important studies of the past five years. No chapter can be considered a comprehensive review of an area.

In this chapter trends in research on the learning and teaching of mathematics are discussed following a brief summary of the research in areas 5-8 listed at the beginning of the chapter. Treatment of these four areas is brief because the quality of recent research was not high enough

to warrant separate discussion, the area overlapped substantially with other areas, or recent excellent reviews were published elsewhere.

### *Summary of Studies Related to the Effectiveness of Instructional Programs*

The products of the curriculum reform movement made the most obvious contribution to changing educational practices. New texts, materials, instructional procedures, and even organizations for instruction flooded the educational marketplace. However, new materials or procedures of high quality are not insured by their being developed by well-funded scholars. Development is an engineering process designed to produce products. The only way to establish quality is by systematic evaluation.

Unfortunately, the science of curriculum evaluation is in its infancy, and little of the vast resources available in the last decade has been used to nurture it (although some evaluation strategy was followed by most developers). Student performance on some test at the end of instruction was usually employed to assess effectiveness. Neither ongoing effectiveness nor the utility of what was learned was examined systematically; evaluative judgments were often based on a single terminal measure. A stereotype of most evaluations would include a description of the data gathered and a comparison of that data with data from a nonequivalent group.

#### *Descriptive Studies*

In the scientifically weakest but perhaps most practical evaluations, innovators such as Davis (1905) and Johtnz (1967) admitted the difficulties of systematic determination of effectiveness. Davis (1965) simply described the diversity of classroom activity, the amount of active student participation, and the creative classroom experiences as evidence to support the effectiveness of the third and fourth years of the Madison Project. Johtnz (1967) examined the effect of a trained mathematician teaching disadvantaged elementary school children. He expected the scholar to communicate advanced notions of mathematics to the children and to arouse their interest in schoolwork. Observers' comments were used to describe the effectiveness of Project SEED (Special Elementary Education for the Disadvantaged). Thus, evidence other than student achievement can and should be used to examine curricular effectiveness. It is unfortunate that systematic procedures for gathering and reporting this kind of evidence were not used.

Reports of some studies such as Shah's (1969) only contained test results. No attempt was made to relate the results obtained to other data. Shah taught children of ages 7 to 11 a carefully developed set of geometry



lessons. He administered a content test and reported generally high scores; the assumption was that without the lessons, low scores could be expected. Another common evaluation procedure is to report only gain scores. Pollack (1967) reported the results of a study on the use of the School Mathematics Study Group (SMSG) set of films and accompanying text to retrain elementary school teachers. Although a variety of instructional procedures was used, generally uniform and large gain scores were reported for scores derived from nominally parallel tests.

If descriptive procedures alone are to be useful, they must be comprehensive. Cronbach (1963) challenged developers to gather descriptive data during development by systematic procedures, subsequently called formative evaluations of the product. Exploratory studies by Coulson (1964) and Romberg and Roweton (1969) reported the use of formative procedures in developing instructional programs. Coulson began with an instructional program on geometric inequalities and used information gathered in tutorial sessions to modify the materials. To make a number of revisions in materials and procedures, Romberg and Roweton periodically gathered data on an instructional program being developed for kindergarten and first grade students.

### *Comparative Studies*

Comparative evaluations are more common than strictly descriptive ones. Typically comparisons are made on data from experimental and control groups. However, true control groups are rare; they are usually alternative treatment groups. Studies by Biddle (1967) and Berger and Howitz (1967) were typical of group comparison studies. Two groups were given the same opportunity to learn a course of study, one group learned in a traditional way and the other in an experimental way. In Biddle's study the course was programmed geometry, and in Berger and Howitz's study the course was a discovery program for general mathematics students. Although results of such studies differ on some details, the usual finding is "no significant differences." Occasionally differences between groups are found; Biddle found that, although there were no significant differences in achievement, the "regular" class had better retention.

Some studies in this category compare effects of "modern" and "conventional" programs over longer periods of time. Hungerman (1967) and Grafft and Ruddel (1968) examined students at the end of Grade 6 who had been in either the SMSG curriculum or a conventional program during Grades 4, 5, and 6. Hungerman found that the SMSG students were superior on a contemporary test and the non-SMSG students were superior on a conventional test. Grafft and Ruddel reported large differences favoring the SMSG students on three measures: understanding,

thought processes, and a transfer test. No differences were reported on a computation test.

In other studies, such as those by Suppes and Binford (1964) and Herriot (1967), the relative effectiveness of a program in quite different populations of subjects was evaluated. Suppes and Binford concluded that bright fifth graders could learn and use symbolic logic as well as college students. The treatment time was approximately the same for both groups but considerably distributed for the fifth graders. In an exploratory study, Herriot examined the differences between regular seventh and ninth grade students taking a one-year SMSG course and students classified as "slow learners" who were enrolled in two-year courses covering basically the same material. Despite variation in the way in which schools classify students as "slow learners," such students clearly did benefit from the slower pace of instruction, although they failed to achieve as well as the regular students.

Scriven (1967), in replying to Cronbach's challenge to conduct formative evaluations, pointed out the need for summative evaluations. Such procedures call for multivariate comparisons among alternative programs, a sort of consumer report for curricula. The initial reports (School Mathematics Study Group, 1969) from the SMSG-directed National Longitudinal Study of Mathematical Ability (NLSMA) promise to be a prototype for demonstrating comparisons on a large number of measures between several groups defined by the text series used. Reports will be made for Grades 4 through 12. The data from NLSMA promises to be a major stimulus to research in mathematics education, as the SMSG texts have provided stimulation to mathematics instruction in the past decade.

### *Experimental Studies*

In their designs for comparative studies on instructional programs, researchers attempt to control such factors as teachers or procedures. Thus, experimental studies are those in which researchers attempt to account or control for the variations in performance that might be attributable to other factors in the instructional situation. Eriksen and Ryan (1966) controlled for the teacher effect when comparing "modern" and "conventional" programs by having each teacher teach both programs. The most ambitious experimental study was done by Armstrong (1968). She experimentally controlled for teacher, instructional method, curriculum organization, classroom environment, media, and five learner variables. By determining weights associated with each variate in a series of canonical correlations, Armstrong demonstrated that the investigator who uses a global measure to assess experimental effects may obtain no significant differences because of cancelling interaction effects.

### *Other Studies*

Curricula are not the only products of the reform movement in mathematics. Teaching procedures (particularly discovery teaching), manipulation aids, and even organizations for instruction have been produced. Studies of the effectiveness of teaching procedures and manipulative aids are reviewed by Fey and Kieren in this issue. The effectiveness of various organizational procedures for the teaching of mathematics at the elementary level was last reviewed by Weaver (1966), whose summary contained 92 references.

Although results of studies conducted since that time vary, organizational patterns in general appear to make little difference on mathematics achievement scores. However, in most studies both teachers and pupils preferred the alternative organizational patterns to the traditional age-grade self-contained class.

Although it has long been assumed that studying mathematics is useful (if not necessary) for many activities, very few studies have been done to validate this assumption. Not a single study on the utility of mathematics to learning tasks outside mathematics was reported. Grafft and Ruddel (1968) and Twelker (1967) examined the usefulness of mathematics to the later learning of mathematics. In a rather loosely designed comparative study. Grafft and Ruddel found that students who had studied MSG in grades 4, 5, and 6 were able to learn a new modern topic in mathematics more rapidly and effectively than were students who had been in a traditional program. Twelker carefully designed an experiment on types of teacher-learner interaction in learning by discovery. Although he found no differences in learning a new task which could be attributed to type of prior learning, he demonstrated a methodological procedure that could be used with more contrasting treatments.

*Additional References:* Concannon (1966); Klinkerman and Bridges (1967); Lewis (1966); Price, Prescott, and Hopkins (1967).

### *Summary*

Studies on the effectiveness of instructional programs, although numerous, were of poor quality. The major reason is that most efforts were expended in developing products. Researchers were primarily interested in the quality of the mathematics or the ingenuity of the technique and not in the changes in behavior produced. It may be that creative temperament of a developer is not compatible with the objectivity needed to evaluate a program or that program quality was affected by poor conceptualization, lack of adequate funds or other resources, inadequate evaluation instruments, or inappropriate use of experimental procedures and inferential statistics.

However, changes in the quality of evaluations seem to be imminent. There appears to be a growing awareness of the necessity for systematic evaluations and the unique problems they present. It is true that such educational spokesmen as Cronbach (1963), Scriven (1967), Stake (1967), and Wittrock (1966) have challenged developers to evaluate and offered them substantive procedural suggestions. However, since these scholars were not involved in developing mathematical materials, they were not heard. Only recently has the mathematics community been listening. Two documents in press, *Disciplined Inquiry for Education and Recommendations for Curriculum and Instruction Materials*, should help to improve evaluations. The first, edited by Cronbach and Suppes (See American Educational Research Association, 1969) was prepared by the Committee on Educational Research of the prestigious National Academy of Education. The authors of this report argued that decision-oriented studies (evaluations are one kind of decision-oriented studies) have been unsatisfactory, and they spelled out the unique functions and conditions of such studies. The second document, *Recommendations*, was written by Tyler and Klein (1967) for the Curriculum Evaluation Committee of AERA. This report was designed as a set of guidelines, recommendations, and technical standards for evaluating curricula. It was deliberately modeled after the influential *Standards for Educational and Psychological Tests and Manuals* (Joint Committee of the American Psychological Assoc., American Educational Research Assoc., and National Council on Measurement in Education, 1966).

One can safely predict that when the *Review* again looks at mathematics education a large number of well-conceived systematic evaluations will have been carried out.

### *Studies Relating Learner Aptitudes and Abilities to Mathematics Learning*

The ugly iron curtain described by Stolurow (1965), which divides the psychological area of individual differences from the area of learning, still exists. Stolurow's scholarly review of this problem as it relates to mathematics is the most cogent analysis of the potential contribution the areas can make to each other and to the teaching of mathematics (also see Walbesser, 1965). Unfortunately, this valid analysis appears in an obscure conference report which undoubtedly has not been read by most mathematics educators. The studies in this area are reviewed here in two categories: the identification of mathematical abilities, and the development of mathematical abilities.

#### *Mathematical Abilities*

In contrast to the large volume of curricular studies, only a few



studies have been done recently on identifying mathematical abilities. The classic studies by the Swedish psychologist Werdelin on mathematical ability (1958) and geometrical ability (1961) are still little known in this country. The promise in NLSMA's title (National Longitudinal Study of Mathematical Abilities) appears to be unfulfilled since all the initial reports (School Mathematics Study Group, 1969) are curricular comparisons.

Correlational techniques have been used to examine some factors affecting ability. Hedley (1968) studied the relationship of personality factors and mathematical ability. She found two canonical factors, one which contrasted computation and clerical skills with avoidance tendencies, femininity, and social service. The second contrasted high general intelligence and low computation with high sensitivity and self-doubt. Westbrook et al. (1965) and Leton and Kim (1966) correlated various intellectual abilities and mathematics achievement. Factor analyses (at grades 4, 5, and 6 in the Westbrook et al. study and at grade 9 in the Leton and Kim study) revealed numerical reasoning and ability to discern verbal meaning to be highly correlated with achievement.

Williams (1965) and Leiderman, Chinn, and Dunkley (1966) described the mathematical abilities of primary-school children. Williams worked with children entering kindergarten and Leiderman et al. studied culturally disadvantaged and advantaged kindergarten and first grade children, but in both investigations the variability of the populations on a variety of tasks related to mathematics was striking. Leiderman et al. found significant differences favoring advantaged over disadvantaged children on four of eleven measures at kindergarten and seven of eleven at grade 1.

A number of investigators continue to study the relationship between mathematics and reading ability. Since mathematics has its own symbolism and syntactics, it requires its own reading skills. Several investigators (e.g., Smith and Heddons, 1964) employed readability formulas to analyze mathematics texts. Others (e.g., Kane and Hater, 1968) tried to adapt standard reading techniques to the readability of mathematical English. Call and Wiggin (1966) demonstrated that a ten-day unit on the reading of mathematics helped students to solve work problems. Surprisingly, Gilmery (1967) found that remedial reading instruction had a positive effect on arithmetical computation achievement.

### *The Development of Mathematical Abilities*

One of the phenomena of the past decade has been the intellectual stimulation and volume of research studies generated by the observations and theories of Jean Piaget. Piaget's vast conceptualization of human cognitive development is of particular import to mathematics educators

since most of his observations have been on mathematical tasks (quantification, geometry, spatial relations, logic, etc.).

A review of all the Piagetian studies that have been done is not possible in this chapter. Fortunately, two Canadian scholars recently prepared excellent critical reviews from quite different points of view. Sullivan (1967) attempted to draw the implications of Piagetian theories for educational practice. He concluded that the contributions are more apparent than real because of uncritical extrapolation of Piaget's observations; clearly this is a fault not of Piaget but of his followers. Harrison (1969) carefully reviewed 80 studies and attempted an extrapolation of the findings to mathematics learning and instruction. He expressed optimism about the relevance of much of Piaget's work to instruction.

*Additional References:* Carroll (1963); Cronbach (1957); Klausmeier et al. (1969); Mehler and Bever (1967, 1968); Nuffield Mathematics Project (1967, 1968); Piaget (1968); Ripple and Rockcastle (1964).

### *Summary*

It seems plausible that many instructional procedures in mathematics could be guided by appropriate utilization of information on cognitive individual differences, but this is not yet the case. As individualization of instruction and computer management become a reality, aptitude and ability data should become extremely useful. Perhaps in the next decade, cracks in the iron curtain will appear.

### *Studies on Attitudes Toward Mathematics*

Although conceptual and methodological problems plague many areas of inquiry, they are particularly acute in the study of attitudes toward mathematics. One of the anticipated outcomes of the modern mathematical programs was that students' attitudes toward mathematics would greatly improve. Although many investigators studied attitudes toward mathematics, evidence to support the claim of improved attitudes is still meager a decade after the new programs were introduced.

Attitude studies have not been fruitful for many reasons. First, attitudes have been operationally defined from scores on paper-and-pencil tests which are beset with insurmountable validity, internal consistency, and score stability problems. A theoretical formulation is needed which conceives of attitudes as a set of moderator variables that affect the subject's response to mathematical situations in observable and predictable ways. Second, most investigators use a single, global measure of attitudes toward mathematics. This is certainly not realistic, since there is probably a set of predispositions or feelings that vary from computation to problem

solving, etc. Third, even if researchers knew what attitudes were and could measure them, they would need to identify the procedures that might be used to modify existing attitudes. There is no reason to believe that the new mathematics programs were appropriately designed to modify the underlying feelings of students or teachers.

Despite these problems, considerable work was done in the last decade on attitudes toward mathematics. Aiken (1969) prepared a careful, critical review of 99 of these studies.

### *Studies Related to Evaluation and Measurement of Mathematics Achievement*

It is safe to generalize that, in most mathematics studies conducted during the 1960's, researchers used inappropriate or inadequate measuring devices to assess mathematics achievements. Since there were changes in content, objectives, and even in the purpose of gathering data, new tests were badly needed. However, most investigators continued to use tests developed earlier—tests designed to measure other objectives and constructed to maximize individual differences for counseling purposes rather than to measure the new objectives and to evaluate the effectiveness of instructional programs.

The time lag between changes in instructional programs and the development of appropriate tests is understandable, if unfortunate. Program developers were primarily interested in making the stimulus or content characteristics of mathematics valid, but not in describing the behavioral changes such content would produce. Most mathematicians were satisfied with describing their aims or goals as general declarations of intent; however, it is essential to convert their inspirational phrases into detailed descriptions of behavior that can be accepted as evidence of learning. Considerable work was done recently on specifying behavioral objectives. The belief that global, individual difference tests are inadequate sources of data for making program decisions about groups has led to ideas of mastery learning and criterion-referenced tests rather than comprehensive achievement and norm-referenced tests.

#### *Objectives*

The translation of goals into behavioral objectives has long been advocated by evaluators; the curriculum revolution in mathematics made that translation a necessity. Concurrent with the revolution the *Taxonomy of Educational Objectives: Handbook 1, Cognitive Domain*, (Bloom, 1956) was published. This useful classification of objectives proved to be widely followed in mathematics. Wood (1967), in summarizing the recent efforts of many mathematical groups to specify objectives, found that each owed

considerable debt to the *Taxonomy*. Such influential mathematical groups as the College Entrance Examination Board (1960), SMSG-NLSMA (Romberg and Wilson, 1968), and the group conducting the International Study of Mathematics Achievement (Husén, 1967) adopted classifications of objectives similar to that outlined in the *Taxonomy*. These classifications have been used to construct more appropriate comprehensive mathematics tests (Epstein, 1968), to critique existing or newly developed tests (Romberg, 1967, 1968), to be one axis of a content by process matrix for the development of tests for specific behaviors (Romberg and Wilson, 1969), and to be a classification device for the development of an item bank (Wood, 1967).

Avital (1968) and Romberg (1967) attempted to validate taxonomic classifications with reference to mathematics learning at Grade 9. Both got high agreement among judges on item categorization and demonstrated a quasi-simplex ordering of the taxonomic categories; however, both investigators had some problem in interpreting achievement data by taxonomic levels. Romberg found that subscales derived by factor analytic techniques failed to have any taxonomic or even content characteristics; instead, they tended to be grouped on a familiar-unfamiliar continuum. This finding was supported by Pruzek (1967); he compared mathematicians' categorizations of items (content areas) from the Scholastic Achievement Test with factorially derived categories and found no similarity between categorizations. Perhaps achievement is most accurately described by the opportunity to learn and not by objectives. Although these studies may cast some doubt on the validity underlying taxonomic categorization, the utility of the taxonomic approach is unquestionable.

A more specific procedure for specifying behavioral objectives, called *task analysis*, has long been advocated by several men (Glaser and Klaus, 1962; Gagné, 1967; Walbesser, 1968). However, with the exception of a few specific units for experimental studies, no comprehensive task analysis has been completed in mathematics. The study by Gagné and Paradise (1961) is a good example of the use of task analysis. Although it was not derived by task analysis, the most comprehensive set of objectives for a particular instrumental program is that associated with the Individually Prescribed Instruction (IPI) mathematics program for the elementary school prepared by the University of Pittsburgh Learning Research and Development Center (1967).

### *Mastery Learning and Criterion-Referenced Tests*

Comprehensive tests given at the end of a course to compare experimental and traditional programs are almost useless for making practical decisions, so critical to development, about whether students have acquired the specific concept or skill taught in a lesson. Rather, precise



evidence is needed at the time it is taught; specific information about expectations and not just comparing groups is required. Such learning has been called *mastery learning* (Bloom, 1968), and such tests *criterion-referenced tests* (Glaser and Klaus, 1962). Several studies associated with the IPI program (Cox and Graham, 1967; Cox and Vargas, 1966; Wang and Lindvall, 1969) were investigations of the utility of CRT's for mastery learning as well as of some of the methodological problems (like item selection) associated with the procedure. One of the problems with CRT's, namely appropriate sampling of test items from a well-specified domain, was examined by Hively, Patterson, and Page (1969). The area of mastery learning and criterion-referenced tests is just now becoming of interest to many measurement specialists. In the next few years substantial work will undoubtedly be done on this area. Additionally, prospects look bright for the development of achievement tests reflecting changes in mathematics education as those changes are made.

*Additional References:* Cahen, Romberg, and Zwirner (1968); Gorth (1969); Gray (1966); Heimer (1966); Walbesser and Carter (1969).

### *Other Areas of Research*

Like all categorization schemes, the eight categories listed at the beginning of the chapter will not encompass all studies. In particular, comparative education studies and programatic research, now receiving greater emphasis, do not fit the scheme. The efforts of some organizations—the School Mathematics Study Group, the Institute for Mathematical Studies in the Social Sciences, and the U. S. Office of Education Research and Development Centers at Wisconsin or Pittsburgh—are not adequately described by a simple listing of some of the reports in various categories.

#### *Comparative Education*

One of the most widely discussed cross-cultural studies conducted in the last few years was the International Study of Achievement in Mathematics (Husén, 1967). The study is a true comparative educational study, with mathematics achievement as the dependent variable. School children in Australia, Belgium, England, Finland, France, Israel, Japan, the Netherlands, Scotland, Sweden, the United States, and West Germany were included in the study. Data were collected in 1964 by means of a series of multiple-choice tests especially devised for international use. The tests were administered in each country to samples of thirteen-year-olds and pupils finishing the secondary school. Considering that more than 130,000 students, 13,000 teachers, and 5,000 schools from 12 countries were included, this project ranks as a major technical achievement and clearly demonstrates the feasibility of a large-scale international study of educa-

tional problems; however, the educational value of the findings is not as clear. The goal of the project was to identify social and educational practices that influence students' achievement in mathematics, information of value for educational planning and curriculum development. Extreme care was taken to include a representative sample of schools from each country. Although the authors recognized that the study was correlational and did not allow causal relationships to be identified, the discussions and conclusions frequently betray an uncritical acceptance of certain causal interpretations. The most interesting finding was that students' opportunity to learn the test materials (as judged by their teachers) correlated very highly (.95) with the between-country differences in mean achievement, although the magnitude of the observed differences in performance among students in different countries was considerable. If the differences were primarily due to the opportunity to learn the ideas included on the test, then the differences among countries are only a function of the differences in the curriculum and not in other types of social or educational practices. Although the results are somewhat clouded by the extreme redundancy in reporting results and the possibly inadequate procedures for analyzing the data, the two-volume report describes a major international accomplishment.

A second but quite different interesting international project is the new series of translations of Russian works on research in mathematics education (Kilpatrick and Wirsup, 1969). The first two of the fifteen volumes which are in preparation are currently available. Volume I includes an introductory paper by N. A. Menchinskaya, "Fifty Years of Soviet Instructional Psychology," which is a fairly recent (1967) general survey of the literature. The books will contain one or more articles under general headings such as the Learning of Mathematical Concepts, the Structure of Mathematical Ability, and Problem Solving in Geometry. These translations should provide valuable information to researchers about significant work being done in a different culture.

### *Trends in Research in Mathematics Education*

Speculations about what will happen in the future are always difficult. However, by projecting current work and considering the pronouncements and concerns of leaders in the field, certain trends become apparent. First, the largest percentage of recent studies are Ph.D. dissertations, and the number of these being done each year is increasing. In 1964, 75 doctoral dissertations in mathematics education were reported in *Dissertation Abstracts* (Summers, 1967); Mangrum and Morris (1969) found 97 theses in 1966; and Weaver (1969) reported 157 theses for 1968—an increase over 1964 of more than 100%. Two of the serious problems in the field are that few of the theses have been published and that few

researchers have done work beyond their doctoral study. Mathematics education can be a mature field only if its research is widely disseminated and investigators continue to do research. Theses provide most students with their first chance to do research. The quality of such research is often not high, but doing a thesis is a learning experience. The most talented students should continue to do research. Unfortunately, the types of position available to people in mathematics education militate against continued research. Teacher preparation, curriculum development, and service to local schools are rewarded, but research is not. Fortunately, this situation is beginning to change. Leaders in the field are pointing to the growing awareness of the need for research. For example, three recent conferences dealt specifically with needed research. Fehr (1966) summarized a conference held in 1965 which dealt with unsolved problems and with the methods, resources, and means to investigate these problems. The National Council of Teachers of Mathematics and the University of Georgia (Hooton, 1967) sponsored a more ambitious conference dealing specifically with learning, teaching, and curricular problems. Long (1968) summarized a conference held by mathematicians at Cornell University who dealt primarily with the question of what should constitute a doctoral thesis in mathematics education for people majoring in mathematics. Outside of such conferences, a number of influential members of the mathematics education community have spoken out rather strongly on the need for research (E. G. Begle, 1968; Patrick Suppes, 1967; and Joseph Scandura, 1967).

There also were efforts to coordinate work in the field. The National Council of Teachers of Mathematics (NCTM) Research Advisory Committee sponsored a research monograph, cosponsored the Georgia Conference on Needed Research in Mathematics Education, established research sections at its annual meetings, and most recently sponsored a new journal designed specifically for researchers in the fields of mathematical education (see below). Within the AERA, a Special Interest Group in Mathematics Education was formed in 1968 and sponsored paper reading sessions and a symposium at the 1969 AERA Annual Meeting.

One outcome of the recent interest in research in mathematics education is the appearance of new publications for research reporting. In the past *The Mathematics Teacher* and *The Arithmetic Teacher*, prepared primarily for classroom teachers by the NCTM, published the largest percentage of articles dealing specifically with mathematics education. The first issue of NCTM's *Journal for Research in Mathematics Education* (David C. Johnson, editor), mentioned above, appeared in the Fall of 1969. *Investigations in Mathematics Education* from SMSG, edited by J. Fred Weaver, will provide abstracts of major published research on mathematics education in the United States. *Educational Studies in Mathe-*

matics, edited by Henry Freudenthal, is a new international journal published in Holland for educational leaders rather than teachers. The *Journal of Structural Learning*, first published in June 1969, is not solely for mathematics education, although its editor, Z. P. Dienes, and many of its editorial staff are members of the mathematics community. By giving more visibility to research results, these new publications should help to keep mathematicians in the research field.

Four trends in the progress of mathematics education research are apparent. First, there will be more programatic research. Concerted efforts will be made by researchers working in a setting which provides both time and resources for investigating problems from a long-range perspective. Second, there will be more and better basic research. Much programatic research will deal with basic problems about human acquisition of concepts and skills. Third, as stated earlier there will be better curriculum evaluations. Fourth, there will be better tests developed for use in mathematics education.

## References

- Aiken, Lewis R., Jr. *Attitudes Toward Mathematics: A Decade of Research Reviewed*. Athens: Univ. of Ga., 1969. (Mimeo.)
- American Education Research Assoc. *Disciplined Inquiry for Education? Educational Researcher* 20, No. 4: 1, 12; 1969.
- Armstrong, Jenny R. *The Relative Effects of Two Forms of Spiral Curriculum Organization and Two Modes of Presentation on Mathematical Learning*. Doctor's thesis. Madison: Univ of Wis., 1968. Abstract: *Dissertation Abstracts* 29: 141A; 1968.
- Avital, Shmuel M. *A Cognitive Analysis and Validation of the Taxonomy of Educational Objectives with Reference to Mathematics Learning*. Doctor's thesis. Toronto: Ontario Institute for Studies in Education, 1968.
- Begle, E. G. *Curriculum Research in Mathematics. Research and Development Toward the Improvement of Education*. (Edited by Herbert J. Klausmeier and George T. O'Hearn.) Madison, Wis.: Dembar Educational Research Services, 1968. Ch. 5, pp. 44-48.
- Berger, Emil J. and Howitz, Thomas A. *Evaluation of Experiences in Mathematical Discovery. Research in Mathematics Education*. (Edited by Joseph M. Scandura.) Washington, D. C.: National Council of Teachers of Mathematics, 1967. Ch. 6, pp. 60-69.
- Biddle, John C. *Effectiveness of Two Methods of Instruction of High School Geometry on Achievement and Retention: A Two and One-Half Year Study. School Science and Mathematics* 67: 689-94; 1967.
- Bloom, Benjamin S. (editor). *Taxonomy of Educational Objectives. Handbook I: Cognitive Domain*. New York: David McKay, 1956.
- Bloom, Benjamin. *Mastery Learning. Evaluation Comment 1, No. 2*. Los Angeles: Center for the Study of Evaluation of Instructional Programs, Univ. of Calif., 1968.
- Bourne, Lyle E., Jr. *Human Conceptual Behavior*. Boston: Allyn and Bacon, 1966.
- Bruner, Jerome S. *The Process of Education*. New York: Vintage Books, 1960.
- Call, Russell J. and Wiggin, Neal A. *Reading and Mathematics. Mathematics Teacher* 59: 149-57; 1966.
- College Entrance Examination Board. *A Description of the College Board Achievement Tests*. Princeton, N. J.: The Board, 1960.
- Coulson, John E. *Research with a Program on Geometric Inequalities*. Report No. TM-895/102/00. Santa Monica, Calif: System Development Corporation, 1964.
- Cox, Richard C. and Graham, Glenn T. *The Development of a Sequentially Scaled Achievement Test. Journal of Educational Measurement* 3: 147-50; 1967.



- Cox, Richard C. and Vargas, Julie S. *A Comparison of Item Selection Techniques for Norm-Referenced and Criterion-Referenced Tests*. Reprint No. 7. Pittsburgh: Learning Research and Development Center, Univ. of Pittsburgh, 1966. (ERIC: ED 010 517.)
- Cronbach, Lee J. Evaluation for Course Improvement. *Teachers College Record* 64: 672-83; 1963. Reprinted in *New Curricula*. (Edited by Robert W. Heath.) New York: Harper and Row, 1964. Ch. 12, pp. 231-48.
- Cronbach, Lee J. Issues Current in Educational Psychology. *Mathematical Learning*. Monographs for the Society for Research in Child Development, Vol. 30, No. 1. (Edited by Lloyd N. Morrisett and John Visonhaler.) Chicago: Univ. of Chicago Press, 1965. Chapter 4, pp. 109-26.
- Davis, Robert B. (editor). *A Modern Mathematics Program as It Pertains to the Inter-relationship of Mathematical Content, Teaching Methods and Classroom Atmosphere* (The Madison Project). Cooperative Research Project No. D-093. Webster's Grove, Mo.: Webster College, 1965. (ERIC: ED 003 371.)
- Epstein, Marion G. Testing in Mathematics: Why? What? How? *Arithmetic Teacher* 15: 311-19; 1968.
- Ericksen, Gerald L. and Ryan, James J. *A Study of the Effects of Experimental Programs on Pupil Achievement Observed During First Three Years of the Project, Secondary Mathematics Evaluation Project*. Minn. State Department of Education Report No. MNL-TR-66-4. St. Paul: The Department, 1966. (ERIC: ED 011 977.)
- Fehr, Howard A. (editor). *Needed Research in Mathematical Education*. New York: Teachers College Press, 1966.
- Gagné, Robert M. The Acquisition of Knowledge. *Research in Mathematics Education*. (Edited by Joseph M. Scandura.) Washington, D. C.: National Council of Mathematics, 1967. Ch. 2, pp. 6-20.
- Gagné, Robert M. and Paradise, Noel E. *Abilities and Learning Sets in Knowledge Acquisition*. Psychological Monographs, General and Applied: Vol. 75, No. 14 (Whole No. 518). Washington, D. C.: American Psychological Assoc., 1961.
- Gilmary, Sister. Transfer Effects of Reading Remediation to Arithmetic Computation When Intelligence Is Controlled and All Other School Factors Are Eliminated. *Arithmetic Teacher* 14: 17-20; 1967.
- Glaser, Robert. Concept Learning and Concept Teaching. *Learning Research and School Subjects*. (Edited by Robert M. Gagné and William J. Gephart.) Eighth Annual Phi Delta Kappa Symposium on Educational Research. Itasca, Ill.: F. E. Peacock, 1968. Ch. 1, pp. 1-36.
- Glaser, Robert and Klaus, David J. Proficiency Measurement: Assessing Human Performance. *Psychological Principles in System Development*. (Edited by Robert M. Gagné.) New York: Holt, Rinehart and Winston, 1962. Ch. 12, pp. 419-74.
- Glennon, Vincent J. and Callahan, Leroy G. *Elementary School Mathematics: A Guide to Current Research*. Washington, D. C.: NEA Association for Supervision and Curriculum Development, 1968.
- Graffit, William D. and Ruddel, Arden K. Cognitive Outcomes of the SMSG Mathematics Program in Grades 4, 5, and 6. *Arithmetic Teacher* 15: 161-65; 1968.
- Harrison, D. Bruce. *Piagetian Studies and Mathematics Learning and Instruction*. Paper presented at the 47th Annual Meeting of the National Council of Teachers of Mathematics, Minneapolis, April, 1969. Calgary: Univ. of Alberta.
- Heath, Robert W. (editor). *New Curricula*. New York: Harper and Row, 1964.
- Hedley, Carolyn N. The Relationship of Personality Factors to Scientific and Mathematical Ability Factors. *School Science and Mathematics* 68: 265-71; 1968.
- Herriot, Sarah T. *The Slow Learner Project: The Secondary School "Slow Learner" in Mathematics*. SMSG Report No. 5. Stanford: School Mathematics Study Group, Stanford Univ., 1967.
- Hively, Wells; Patterson, Harry L.; and Page, Sarah H. A 'Universe-Defined' System of Arithmetic Achievement Tests. *Journal of Educational Measurement* 5: 275-90; 1969.
- Hooten, Joseph R., Jr. (editor.) Proceedings of National Conference on Needed Research in Mathematics Education. *Journal of Research and Development in Education* 1, No. 1: 1-142; 1967.
- Hungerman, Ann D. Achievement and Attitude of Sixth-Grade Pupils in Conventional and Contemporary Mathematics Programs. *Arithmetic Teacher* 14: 30-39; 1967.

- Husén, Torsten (editor). *International Study of Achievement in Mathematics*. 2 Vols. New York: Wiley 1967.
- Johntz, William F. Innovation and the New Concern for the Disadvantaged. *California Teachers Association Journal* 63: 5-6, 32; 1967.
- Joint Committee of the American Psychological Assoc., American Educational Research Assoc., and National Council on Measurement in Education. *Standards for Educational and Psychological Tests and Manuals*. Washington, D. C.: American Psychological Assoc., 1966.
- Kane, Robert and Hater, Sn. Mary Ann. *Adapting the Cloze Technique to Study the Readability of Mathematical English*. Paper read at Annual Meeting of the National Council of Teachers of Mathematics, Philadelphia, 1968. Lafayette, Ind.: Purdue Univ.
- Kilpatrick, Jeremy and Wirsup, Izaak (editors). *Soviet Studies in the Psychology of Learning and Teaching Mathematics*. 15 Vols. School Mathematics Study Group and Survey of Recent East European Mathematical Literature. Stanford: School Mathematics Study Group, Stanford Univ., 1969.
- Leiderman, Gloria.; Chinn, William G.; and Dunkley, Mervyn E. *The Special Curriculum Project: Pilot Program on Mathematics Learning of Culturally Disadvantaged Primary School Children*. SMSG Report No. 2. Stanford: School Mathematics Study Group, Stanford Univ., 1966.
- Leton, Donald A. and Kim, Sharon. *Analysis of Mathematical Abilities Required for Success in Ninth-Grade Mathematics*. Cooperative Research Project No. CRP-S-343. Honolulu: Education Research and Development Center, Univ. of Hawaii, 1966. (ERIC: ED 010 420.)
- Long, Richard G. *A Report for the Conferees of the Conference on Mathematics Education*. Ithaca, N. Y.: Cornell Univ. Center for Research in Education, 1968. (Mimeo.)
- Mangrum, Charles and Morris, William. Doctoral Dissertation Research in Science and Mathematics. *School Science and Mathematics* 69: 443-49; 1969.
- Pollack, Helen. *A Film-Film Text Study*. SMSG Report No. 3. Stanford: School Mathematics Study Group, Stanford Univ., 1967.
- Pruzek, Robert M. *A Comparison of Two Methods for Studying Test Items*. Doctor's thesis. Madison: Univ. of Wis., 1967. Abstract: *Dissertation Abstracts* 28: 3035A; 1967.
- Romberg, Thomas A. *Derivation of Subtests Measuring Distinct Mental Processes Within the NLSMA Algebra Achievement Test*. Doctor's thesis. Stanford: Stanford Univ., 1967. Abstract: *Dissertation Abstracts* 29: 419A; 1968.
- Romberg, Thomas A. Contemporary Mathematics Test Series. *Journal of Educational Measurement* 5: 349-51; 1968.
- Romberg, Thomas A. and Roweton, Marilyn. *Pilot Developmental Activities in Elementary Mathematics Conducted at Huegel School, Madison, Wisconsin, 1967-68*. Working Paper No. 24. Madison: Wis. Research and Development Center for Cognitive Learning, Univ. of Wis., 1969.
- Romberg, Thomas A. and Wilson, James W. The Development of Mathematics Achievement Tests of the National Longitudinal Study of Mathematical Abilities. *Mathematics Teacher* 61: 489-95; 1968.
- Romberg, Thomas A. and Wilson, James W. *The Development of Tests for the National Longitudinal Study of Mathematical Abilities*. NLSMA Report Vol. 7. Stanford: School Mathematics Study Group, Stanford Univ., 1969.
- Scandura, Joseph M. (editor). *Research in Mathematics Education: An Overview and a Perspective. Research in Mathematics Education*. Washington, D. C.: National Council of Teachers of Mathematics, 1967. Ch. 11, pp. 115-25.
- School Mathematics Study Group, NLSMA Reports. *Newsletter* No. 30. Stanford: the Group, Stanford Univ., 1969.
- Scriven, Michael. The Methodology of Evaluation. *Perspectives of Curriculum Evaluation*. (Edited by Robert E. Stake.) American Educational Research Assoc. Monograph Series on Curriculum Evaluation No. 1. Chicago: Rand McNally, 1967. Pp. 39-83.
- Shah, Sair Ali. Selected Geometric Concepts Taught to Children Ages Seven to Eleven. *Arithmetic Teacher* 16: 119-28; 1969.

- Shulman, Lee S. and Keislar, Evan R. (editors). *Learning by Discovery: A Critical Appraisal*. Chicago: Rand McNally, 1966.
- Smith, K. J. and Heddons, J. W. Readability of Experimental Mathematics Materials. *Arithmetic Teacher* 11: 391-94; 1964.
- Stake, Robert E. (editor). *Toward a Technology for the Evaluation of Educational Programs. Perspectives of Curriculum Evaluation*. American Educational Research Assoc. Monograph Series on Curriculum Evaluation No. 1. Chicago: Rand McNally, 1967. Pp. 1-12.
- Stolurow, Lawrence M. *Learner Characteristics Which Are Relevant to Consider in the Teaching of Elementary Mathematics. A Project for the Improved Use of Newer Educational Media in Elementary School Mathematics: Report of First Positional Conference*. (Edited by James H. Zant.) Sponsored by the National Council of Teachers of Mathematics. Stillwater: Okla. State Univ., 1965. (Mimeo.)
- Sullivan, Edmund V. *Piaget and the School Curriculum*. Bulletin No. 2. Toronto: Ontario Institute for Studies in Education, 1967.
- Summers, Edward G. Doctoral Dissertation Research in Science and Mathematics Reported for 1964. *School Science and Mathematics* 67: 37-50; 1967.
- Suppes, Patrick. The Case for Information-Oriented (Basic) Research in Mathematics Education. *Research in Mathematics Education*. (Edited by Joseph Scandura.) Washington, D. C.: National Council of Teachers of Mathematics, 1967. Ch. 1, pp. 1-5.
- Suppes, Patrick and Binford, Frederick. *Experimental Teaching of Mathematical Logic in Elementary School*. Cooperative Research Project No. D-055. Stanford Univ., 1964. (ERIC: ED 003 367.)
- Twelker, Paul. *Two Types of Teacher-Learner Interaction in Learning by Discovery*. Office of Education Project No. 5-0580. Monmouth: Ore. State System of Higher Education, Teaching Research Division, 1967.
- Tyler, Louise L. and Klein, M. Frances. *Recommendations for Curriculum and Instructional Materials*. Los Angeles: Univ. of Calif., 1967. (Offset)
- University of Pittsburgh, Learning Research and Development Center. *Mathematics Continuum: Individually Prescribed Instruction*. Pittsburgh: Univ. of Pittsburgh, 1967. P. 25.
- Walbesser, Henry H. *Learner Characteristics Which Are Relevant to Consider in the Teaching of Elementary Mathematics: Supplementary Paper. A Project for the Improved Use of Newer Educational Media in Elementary School Mathematics: Report of First Positional Conference*. (Edited by James H. Zant.) Sponsored by the National Council of Teachers of Mathematics. Stillwater: Okla. State Univ., 1965. (Mimeo.)
- Walbesser, Henry H. *An Evaluation Model and Its Application: Second Report*. Miscellaneous Publication 68-4. Washington, D. C.: American Assoc. for the Advancement of Science, 1968.
- Wang, Margaret and Lindvall, C. M. *An Exploratory Investigation of the Carroll Learning Model and the Bloom Strategy for Mastery Learning*. Paper read at the annual meeting of the American Educational Research Assoc., Los Angeles, Feb. 1969. Pittsburgh: Learning Research and Development Center, Univ. of Pittsburgh.
- Weaver, J. Fred. Differentiated Instruction and School-Class Organization for Mathematical Learning Within the Elementary Grades. *Arithmetic Teacher* 13: 495-506; 1966.
- Weaver, J. Fred. Research on Mathematics Education in Dissertation Abstracts, 1968. *Investigations in Mathematics Education* 1; 1969.
- Werdelin, Ingvar. *The Mathematical Ability: Experimental and Factorial Studies*. Lund, Sweden: C. W. K. Gleerup, 1958.
- Werdelin, Ingvar. *The Geometrical Ability and the Space Factor in Boys and Girls*. Lund, Sweden: Univ. of Lund, 1961.
- Westbrook, Helen R. et al. *Intellectual Processes Related to Mathematics Achievement at Grade Levels 4, 5, and 6*. Cooperative Research Project No. CRP-S-046. Athens: Univ. of Ga., 1965. (ERIC: ED 003 669.)
- Williams, Alfred H. *Mathematical Concepts, Skills, and Abilities of Kindergarten Entrants*. *Arithmetic Teacher* 12: 261-68; 1965.
- Willoughby, Stephen S. *Mathematics. Encyclopedia of Educational Research*. (Edited by Robert L. Ebel.) New York: Macmillan, 1969. Pp. 766-77.



- Wittrock, M. C. *The Experiment in Research on Evaluation of Instruction*. Report No. CSEIP-WP-2. Los Angeles: Univ. of Calif., 1966. (ERIC: ED 012 107.)
- Wood, R. Objectives in the Teaching of Mathematics. *Educational Research* 9: 83-98; 1967.

### Additional References

- Cahen, Leonard S.; Romberg, Thomas A.; and Zwirner, Walter. *The Estimation of Mean Achievement Scores for Schools by the Item Sampling Technique*. Research Bulletin 68-39. Princeton, N. J.: Educational Testing Service, 1968.
- Carroll, John B. A Model of School Learning. *Teachers College Record* 64: 723-33; 1963.
- Concannon, Sister J. *An Experimental Study of the Influence of Individual vs. Group Instruction on Spatial Abilities in Preschool Children*. Cooperative Research Report No. 2885. Chestnut Hill, Mass.: Boston College School of Education, 1966. (ERIC: ED 010 126.)
- Cronbach, Lee J. The Two Disciplines of Scientific Psychology. *American Psychologist* 12: 671-84; 1957.
- Gorth, William P. *Instructional Objectives, Achievement Monitoring, and Learning: What, How, Why, and Where of Comprehensive Achievement Monitoring*. Working Paper No. WP-7. Stanford: School of Education, Project CAM, Stanford Univ., 1969.
- Gray, Roland F. An Approach to Evaluating Arithmetic Understandings. *Arithmetic Teacher* 13: 187-92; 1966.
- Heimer, Ralph T. A Teach-Test Procedure for Obtaining Measures of Mathematical Aptitude. *Mathematical Monthly* 73: 883-89; 1966.
- Klausmeier, Herbert J. et al. *Concept Learning: A Bibliography, 1950-1967*. Technical Report No. 82. Madison: Wis. Research and Development Center for Cognitive Learning, Univ. of Wis., 1969.
- Klinkerman, Ginger and Bridges, Faith. Team Teaching in Geometry. *Mathematics Teacher* 60: 488-92; 1967.
- Lewis, William P. *Demonstration of a Nongraded Plan of an Elementary School, Utilizing Team Teaching and Programed Instruction to Facilitate Learning in Reading and Mathematics*. Troy, Ala.: Troy State College, 1966. (ERIC: ED 010 329.)
- Mehler, J. and Bever, T. G. Cognitive Capacity of Very Young Children. *Science* 158: 141-42; 1967.
- Mehler, J. and Bever, T. C. Reply. *Science* 162: 979-81; 1968.
- Nuffield Mathematics Project. (Teacher's Guides) *Beginnings; Computation and Structure 2; I Do, and I Understand; Mathematics Begins; Pictorial Representation; Shape and Size 2*. Published for the Nuffield Foundation by W & R Chambers and John Murray. New York: John Wiley, 1967.
- Nuffield Mathematics Project. (Teacher's Guides) *Computation and Structure 3; Shape and Size 3*. Published for the Nuffield Foundation by W & R Chambers and John Murray. New York: John Wiley, 1968.
- Piaget, J. Quantification, Conservation, and Nativism. *Science* 162: 976-79; 1968.
- Price, Edette B.; Prescott, Arthur L.; and Hopkins, Kenneth D. Comparative Achievement with Departmentalized and Self-Contained Classroom Organization. *Arithmetic Teacher* 14: 212-15; 1967.
- Ripple, Richard E. and Rockcastle, Verne N. (editors). *Piaget Rediscovered: Report of the Conference on Cognitive Studies and Curriculum Development*. Ithaca, N. Y.: Cornell Univ., 1964.
- Walbesser, Henry H. and Carter, Heather L. *Differences in Group and Individually Administered Tests of the Same Behavior*. Paper read at the meeting of the American Educational Research Assoc., Los Angeles, February, 1969. College Park: Univ. of Md.

### AUTHOR

THOMAS A. ROMBERG Address: University of Wisconsin, Madison, Wisconsin Title: Assistant Professor Age: 36 Degrees: B.S. and M.A., Omaha Univ; Ph.D., Stanford Univ. Specialization: Measurement.





## 8: CONDITIONS OF LEARNING IN MATHEMATICS: SEQUENCE THEORY DEVELOPMENT

Ralph T. Heimer\*

*Pennsylvania State University*

The purposes of this chapter are 1) to identify and describe research and development efforts that appear to be fruitful in the development of adequate sequence theory in mathematics learning and 2) to suggest potentially productive approaches to further the development of fruitful sequence theory.

### *Definition of Terms*

*Conditions of learning* refers to the external conditions that are deliberately established and manipulated by the instructor, author, or programmer. Classes of external conditions of potential importance include sequence, mode of communication, feedback to the student, response mode, and the like.

The *structure* of an object is defined by specifying its elements and their interrelationships. An analysis of structure provides no information about the relations of the object under consideration to other objects. Any description of structure is relative to certain elements (units) which are treated as devoid of structure on the particular occasion of description; on a different occasion and in a different context, the structure of those basic elements or units may be taken into account (Russell, 1962).

In this chapter the structure of *content* (e.g., mathematics course content) is specified. Since *content* is a term used inconsistently in education, the elemental units of analysis of content structure vary greatly. Gagné (1967a) specified *curricular structure* (i.e., content structure) in terms of competencies or behaviors to be attained; Hickey and Newton (1964) gave substance to their analysis-synthesis plane of the "knowledge space" by displaying a tree-graph connecting concepts and subconcepts or conceptual elements; this tree-graph display operates as *content structure*.

The connotation of *content structure* may include curricular or instructional considerations that do not characterize the expressions "structure of a discipline," or "logical structure" of a body of knowledge. Schwab

---

\*Drs. John J. Lottes, Texas Christian University, and John Withall, Pennsylvania State University, assisted Dr. Heimer in the preparation of this chapter.

(1961) identified the two components of the structure of a discipline as 1) a body of concepts which define its subject matter and control its inquiries and 2) as the methodology for achieving the goals of the discipline. The logical structure of mathematics refers to a system of terms, formulation rules, axioms, transformation rules, and theorems. *Logical structure* and *structure of a discipline* overlap greatly in mathematics. *Content structure* accounts for one or both of these, and typically also accounts for instructional considerations.

*Instructional sequence* means the order in which the learner interacts with units of content. In relating instructional sequence to content structure, Hickey and Newton (1964, p. 9) saw the problem on instructional sequencing as weaving a "thread up and down, back and forth, through the knowledge space." Briggs (1968) and Gagné (1967a) proposed instructional sequences that require the pupil to follow a particular kind of path through a content structure. Fundamental to each of the above cited positions are the assumptions that instructional sequence is most fruitfully formulated and evaluated in conjunction with content structure and that instructional sequence specifies the way in which the content structure is traversed by the learner.

### *The Nature and Importance of Sequence Theory Development*

Every effort to construct curricula in mathematics demands decisions about structuring the content and designing and ordering instructional tasks. Unfortunately, at present there are few precise, coherent, and empirically testable sets of hypotheses for identifying and guiding those instructional decisions. There are no well-formulated theories of instruction. In short, there are no adequate teaching algorithms.

Two technical developments have dramatically sharpened the focus on the problems of structure and sequence and their interrelationships: programmed instruction (PI) and computer assisted instruction (CAI). The development of instructional sequences in PI and CAI contexts requires extensive preplanning by the course author. Even digressions from a basic core sequence must be preplanned. Furthermore, the most universal insistence upon precise identification of behavioral objectives and the complementary requirement of adequate evaluation of instruction impose a requirement upon the author that the instructional sequences work. At present, a PI or CAI course author can not avoid making a multitude of ad hoc decisions about structure and sequence. Since decisions can not be ignored or avoided, adequate sources of guidance must be established.

One source of guidance would be the development of adequate *teaching algorithms* for mathematics. Hull (1966, p. 1) informally described an algorithm as "a procedure for solving a problem. Usually one has in mind a class of problems, and the algorithm is a procedure which can be applied

to any member of the class, not to just one particular problem. An important requirement is that the individual steps in an algorithm must be completely unambiguous. Moreover, the algorithm must always produce the solution to the problem, and it must do so in a finite number of steps." Thus, designing a teaching algorithm requires specifying a procedure for constructing an instructional sequence so a given class of objectives will be attained, if some specified set of circumstances prevail (e.g., for a particular type of content structure, learner status, etc.) A final requirement for an adequate teaching algorithm on sequence is that efficiency of an instructional sequence constructed under the algorithm is "guaranteed" under some well-defined conditions; this implies that it is possible to subject the teaching algorithm to empirical test.

### *Bounds of the Review*

Scientific inquiry involves 1) theory development, which includes setting forth variables, relating variables to form hypotheses, and relating hypotheses to form theory and 2) evaluating the theory that has been formulated. Karl Popper (1961) suggested four ways to evaluate a scientific system (theory): judging the logical consistency of the system, judging the logical form of the system, judging the system by comparing it with other systems to determine its potential contribution to the field of study, and judging whether predictions derived from the system are supported by empirical observation.

The formulation and evaluation of theory provide the framework for the selection and examination of inquiries relevant to structure and sequence in mathematics instruction. Consequently, major attempts at such theoretical formulations and their evaluation were given first priority in this chapter. Also included in this chapter are a number of investigations that were judged to introduce variables of potential worth, present some methodological ideas about how to develop sequence theory, or provide evaluative methods of special interest. The selection of studies for this chapter has been further restricted to those that were instructor (programmer, author, investigator) controlled and involved preplanned construction of instructional strategies and curriculum materials.

## *Major Theoretical Formulations*

### *Gagné's Hierarchical Structure and Sequencing*

Gagné's idea of a learning hierarchy and its many ramifications for the analysis and development of instructional actions is important in the study of sequencing. Gagné (1968, p. 5) characterized learning hierarchies as "an ordered set of intellectual skills such that each entity generates a substantial amount of positive transfer to the learning of a not-previously-



acquired higher-order capability." He also wrote that "the learning of each entity also requires the recall of relevant verbalizable knowledge, which, however, is not itself represented in the hierarchy" (p. 6). This last statement represents an important refinement—or at least a clarification—of Gagné's (1962) original ideas about learning hierarchies.

Some of the important assumptions supporting Gagné's (1967b, p. 177) ideas about learning hierarchies are:

1. Any human task may be analyzed into a set of component tasks which are quite distinct from each other in terms of the experimental operations needed to produce them.
2. These task components are mediators of the final task performance, that is, their presence insures positive transfer to a final performance, and their absence reduces such transfer to near zero.
3. The basic principles of training design consist of: (a) identifying the component tasks of a final performance; (b) insuring that each of these component tasks is fully achieved; and (c) arranging the total learning situation in a sequence which will insure optimal mediational effects from one component to another.

These statements certainly imply a set of principles which would have very different names from those we are now most familiar with. They are concerned with such things as *task analysis*, *intra-task transfer*, *component task achievement*, and *sequencing* as important variables in learning and consequently in training. These principles are not set in opposition to the traditional principles of learning, such as reinforcement, differentiation of task elements, familiarity, and so on, and do not deny their relevance, only their *relative importance*.

A critical analysis of the nature and role of learning hierarchies gives rise to a number of serious issues which need to be explored: 1) How is a learning hierarchy constructed and under what conditions is one considered valid? 2) What is the relationship between an hypothesized learning hierarchy and the associated presentation sequence for instruction? 3) What are the "intellectual skills" that make up a learning hierarchy? 4) What is the connection, if any, between the (logical) structure of the content and the design of the associated learning hierarchy?

In response to the first question, Gagné (1968, p. 3) suggested that for each task encountered, one simply ask, "What would the individual already have to know how to do in order to learn this new capability simply by being given verbal instructions?" The validity of the resulting

hierarchy needs to be tested, and Gagné (1967a) and others suggested two testing methods: 1) to check out empirically the individual hypotheses about transfer represented by the hierarchy and 2) to work with the entire hierarchy.

In terms of theoretical considerations about learning hierarchies, Gagné (1963, p. 624) hypothesized that "an individual will not be able to learn a particular topic if he has failed to achieve any of the subordinate topics that support it." This hypothesis was tested in several studies (Gagné and Paradise, 1961; Gagné, 1962; and Gagné et al., 1962); one procedure employed was to administer an examination immediately after the instructional program to establish pass-fail scores for each of the components of the learning hierarchy. Pass-fail patterns were defined that were considered to be either consistent or not consistent with the hypothesis or indicative of a weak instructional program. The ratio of the number of observed patterns consistent with the hypothesis to the total number of observed patterns was formed. In no instance was a ratio of less than .91 reported, which seems to provide some support for the hierarchy hypothesis.

Merrill (1965) reported a study that seemingly offers contrary evidence. He tested the hypothesis that learning and retention of a hierarchical task are facilitated by mastering each successive component of the hierarchy before continuing in the instructional program. Merrill insured mastery by channeling a student who erred on any particular component into a two-stage correction/review procedure. The results of his study seem to indicate that it is not necessary to master one level before proceeding to the next. Briggs (1968) suggested, however, that Merrill's task analysis might be faulty, and this raises a thorny research issue. It does not appear to be logically sound to frame and attempt to test hypotheses about characteristics of learning hierarchies in the absence of establishing bases for deciding when one has a learning hierarchy and when one does not. Otherwise, for every hypothesis the investigator has the option of deciding whether the learning hierarchy was valid or invalid. This, of course, is unacceptable.

About the connection between a learning hierarchy and the associated presentation sequence, Gagné (1968, p. 5) wrote: "I am not sure that a learning hierarchy is supposed to represent a presentation sequence for instruction in an entirely uncomplicated way. Presumably, there should be some relation between an ordered set of intellectual skills and an ordering of a sequence of presentation of a set of frames or topics in an instructional program." Others have not been quite so guarded in their interpretation of the role of the learning hierarchy for developing instructional sequences. Briggs (1968, p. 6), in discussing the learning hierarchy presented in the Gagné and Paradise (1961) article stated that "These subordinate competencies . . . must be taught in a particular sequence

(with some options within layers of the pyramid) rather than in a random sequence. . . ."

Gagné (1968, p. 4) wrote that the "intellectual skills" which make up a learning hierarchy are what some writers call "cognitive strategies. . . . What they are not is just as important. They are not entities of verbalized knowledge. I have found that when deriving them, one must carefully record statements of 'what the individual can do,' and just as carefully avoid statements about 'what the individual knows'."

The connection, if any, between the (logical) structure of the content and the design of the associated learning hierarchy seems to be best characterized at the present time as a "blur." There is most assuredly an intimate relationship between "what an individual can do" and "what an individual knows," but there is a conspicuous absence of information about this relationship in the literature.

It is interesting to note that the learning hierarchy idea—however imprecisely defined and inadequately verified it may be—is already being widely used as a cornerstone for developing instructional sequences in mathematics. In the TRAC (Teaching Research Automated Classroom) procedure described by Kersh (1967), two hierarchies are constructed: a *knowledge* hierarchy which is obtained by asking what the learner must know, and a *complex behavior* hierarchy which is obtained by asking what the learner must do to acquire the knowledge and the complex behavior. The instructional sequence, then, apparently is designed to conform to these hierarchies; that is, the role of these hierarchies in sequencing seems to be taken as axiomatic. The curriculum development efforts of the IPI (Individually Prescribed Instruction) Project apparently also have been heavily influenced by the learning hierarchy idea. Bolvin (1968, p. 238, 239), in discoursing on the implications of the individualization of instruction for curriculum and instructional design, stated that "Curriculum design relates to the determination of behavioral objectives selected on the basis of philosophy of education and the structure of the subject matter under consideration. . . . Precisely stated objectives permit the analysis of the behaviors required as prerequisites to a given objective. This analysis serves as a guide to the curriculum designer in sequencing and ordering the objectives." In connection with the problem of ordering objectives, Bolvin (p. 40) suggested that the technique of *task analysis* is useful in the sense that "it serves as a guide to determining what objectives are useful in attaining another objective." Moreover, he said: "By grouping these related tasks under general characteristics it is possible for the teacher and the student to focus both on the individual objective and the larger and more integrating experience." Once a sequence has been established in conformance with the foregoing ideas, Bolvin asserted that a major question that must be asked is whether the sequence will, for certain

learners, permit variations. Presumably this is to be determined empirically. The impact of Gagné's theory about learning hierarchies on the curriculum construction being done by IPI is unmistakable.

A careful analysis of the literature on learning hierarchies—as described by Gagné—and their role in the development of presentation sequences makes the following conclusions seem tenable: 1) there are no well-defined algorithms for producing learning hierarchies; 2) the connection between the (logical) structure of the knowledge and the associated learning hierarchy has not yet been adequately explored; and 3) the role of learning hierarchies in the development of presentation sequences is unclear.

### *Suppes's Mathematical Models and Sequencing*

Suppes has been another substantial contributor to the development of theoretical constructs with implications for sequencing in mathematics. Of primary importance are his mathematical models that deal with variables potentially important in the design of presentation sequences. Suppes also made important attempts to conceptualize some of the psychological variables that are used and discussed in the literature, but which have resisted adequate definition.

Suppes's mathematical models cover a broad range of phenomena and vary widely in sophistication. There is no question, however, that many of his models fit nicely under the "sequencing" rubric.

Suppes, like Gagné, subscribes to the idea of the importance of accounting for content structure in the study of learning and sequencing. Suppes, Hyman and Jerman (1966, p. 160) stated: "In the cognitive domain mathematics provides one of the clearest examples of complex learning and performance, for the structure of the subject itself provides numerous constraints on any adequate theory." A substantial amount of Suppes's work reflects the attitude contained in the following statement (Suppes, 1966, p. 145):

For anyone interested in the psychological foundations of mathematical concept formation it is natural to ask what is the sort of connection that holds between the logical structure of mathematical concepts and the psychological processes of acquisition of the concept. As far as I know, not very much has been written in the psychological literature about this kind of question. My present view, based partly on our experiments and partly on conjecture, is that the psychological stratification of mathematical concepts will seldom, if ever, do violence to the logical structure of these concepts; but it will markedly deviate from the mathematical analysis of the same concepts with respect to the amount of detail that must be considered.



Suppes reported (1966) details of a study done in collaboration with Rose Ginsburg on the effects of training children on identity of ordered sets before training them on identity of unordered sets. In analyzing the finding that this particular sequence effected negative transfer, Suppes suggested that the result was not surprising since certain principles of interference have been confirmed under conditions of paired-associate learning. In the same discussion he (1966, p. 145) wrote: "The interesting question is: Can we more sharply identify the point of negative transfer in terms of a structure of subconcepts?" The importance of this particular study is that it supports the connection between content structure and sequence which both Suppes and Gagné have hypothesized. Also, it provides evidence that learning tasks can be arranged in a way that gives rise to negative transfer, a situation that an adequate theory of sequencing must consider.

Examples of Suppes's mathematical models that have fairly direct implications for sequencing are his "Process Models for Arithmetic" and "Constant Error Rate Models" (1967). In the case of the former, he constructed a multi-linear regression model which has been adapted to the problem of predicting the observed difficulty exhibited by students on various types of items from a particular class of fourth grade addition problems. Three structural variables are used in the model: MAGSUM—magnitude of the sum; MAGSMALL—magnitude of the smallest addend; and MSTEPS—the number of steps in a very explicit analysis of the algorithm a student would probably use in solving the problem. Data collected to test the model reveal an adequate, though far from perfect, fit.

Suppes did not explicitly indicate his rationale for wanting to make predictions of this sort, but potential applications are obvious in the implementation of the "strands" program he and his colleagues (Suppes, Jerman, and Dow, 1968) developed. Another possible use is in the application of his "Constant Error Rate" models (1967). The rationale underlying these latter models is to organize curriculum materials so the error rate is kept constant as the student learns to deal with more and more difficult items. Such a procedure appears to be at variance with common school practices. However, the "Constant Error Rate" model is under development and has not yet been experimentally validated.

Suppes also designed a multiple-linear regression model for predicting observed response latencies for the same class of addition problems mentioned earlier. Success in prediction of response latency was comparable to that achieved in predicting observed difficulty. Suppes, Hyman and Jerman (1966, p. 163) stated:

From the standpoint of analysis of performance, latencies are in many respects more important as a source of information to the

theorist than as response data. This is particularly true of any studies devoted to skill after a good deal of learning has taken place. As some of the data reported here show and as one would expect anyway on a priori grounds, the range of latencies observed in a group shows systematic variation in a way that clearly reflects a measure of item difficulty. What is ultimately desired in this case is the kind of model that can predict from the structure of an item the process a subject must go through in finding the correct response.

The suggestion seems to be that response latencies may be a good device for determining the nature of the cognitive strategies students use for solving certain classes of problems. This information could have significant implications for sequencing. Suppes and Groen (1966) tested five models of possible strategies used by first graders in solving addition problems of the type  $m + n = \_\_\_$  where  $m + n \leq 5$ ,  $m \geq 0$ , and  $n \geq 0$ . Of the models examined, the one which corresponds to choosing the larger of the two addends,  $m$  and  $n$ , and counting by 1's from that point—using the remaining addend—to reach the desired sum seemed to best fit the data. No claim was made by the authors that this procedure was the one generally employed by the students, but the study does indicate how latencies might be employed to provide fruitful information for the study of sequencing.

### *Other Theoretical Formulations and Research on Sequencing*

#### *Advance Organizers*

Ausubel's position on learning hierarchies and sequencing seems close to Gagné's; Ausubel (1963, p. 86) stated: "Most complex tasks, particularly those that are sequential in nature, can be analyzed into a hierarchy of component learning sets or units. . . . sequential organization of subject matter can be very effective, since each new increment of knowledge serves as an anchoring post for subsequent learning. This presupposes, of course, that the preceding step is always clear, stable and well organized. If it is not, the learning of all subsequent steps is jeopardized. Hence new material in the sequence should never be introduced until all previous steps are thoroughly mastered." It would seem, therefore, that Ausubel views specific topics in a curriculum as advance organizers—of a kind—for subsequent related topics. Ausubel's research with advance organizers is well known and has been widely reported. Merrill and Stolurow (1966) hypothesized that (a) subjects who received a hierarchical presentation (summary prior to test section) would make fewer errors during learning and subsequent testing than would subjects who received a problem-

oriented presentation of the same task (summary after and related to each question) and (b) that subjects who were presented a general review or a specific review when they made errors would make fewer errors during learning and subsequent testing than would subjects who were given only the correct answer. Their results supported the first hypothesis and the specific review part of the second hypothesis. The authors contended that the foregoing findings lent "support to Ausubel's (1963) recommendation for the use of advance organizers and to Gagné's (1965) assumption that hierarchical tasks should be presented level by level."

Ausubel, Robbins and Blake (1957) reported another study involving meaningful school material. They concluded that interpolated learning of material substantively similar, but not identical, to the original learning task does not result in retroactive inhibition; it induces as much retroactive facilitation as identical repetition of the learning material. However, questions about the nature and use of advance organizers in mathematics instruction have been raised by several investigators. Woodward (1966) suggested that Ausubel's criteria for an advance organizer were not well defined, and hence, judgments involved in the construction of particular advance organizers were extremely subjective. Scandura and Wells (1967, p. 295) stated: "By advance organizer, then what Ausubel seems to be referring to is a general non-technical overview or outline in which the non-essentials of the to-be-learned material are ignored. Such introductory material has typically proved easier to learn than the more technical material that follows." They (p. 296) added "Presenting mathematical abstractions, on the other hand, normally involves the use of words or symbols having no referential meaning for the naive student. For this reason, *descriptions* of concrete models of abstract mathematical ideas, although the models themselves involve extraneous features, may be more readily interpretable than formal presentations of the corresponding abstract ideas in terms of the underlying definitions and axioms."

### *Scrambled vs. Ordered Sequence*

The research literature contains a number of studies about the effects of scrambling "ordered" sequences, but the purposes of these studies have not always been clear and their over-all results have not been conclusive. Roe, Case, and Roe (1962) reported a comparative study in which a 71-item program on elementary probability was presented to two groups; one group received the program in its normal ordered form, and one received a scrambled version of it. A criterion test was administered to each student immediately upon completing the program. There were no significant differences reported on time required for learning, error score during learning, criterion test score, or time required for criterion test. In a subsequent study, Roe (1962) reported contrary results with an extended version of

the probability program mentioned earlier in which a random sequence group performed significantly worse on learning time, errors made during learning and on post learning test scores. Roe (1962, p. 409) concluded that "careful sequencing of items has a significant effect on student performance, at least for programs of some length and complexity."

Payne, Krathwohl and Gordon (1967, p. 125) stated: "No one seems to doubt that were one to scramble a whole course that learning would be retarded, so that in part, the size of the unit in which sequence is destroyed is a factor. . . . there may be a continuum of dependence on sequence. At one extreme of the continuum, scrambling may have no effect on learning a set of spelling words which has no logical structure. At the other extreme, scrambling would be expected to result in considerable decrease in learning if the learning of one concept were prerequisite to learning the next in a logical hierarchy." In a study designed to test the foregoing assumption, Payne et al. (1967) examined the effect of scrambling upon the learning of three programs which varied in the judged logical inter-relatedness of the material in each program from low to fairly high. Both immediate and delayed retention tests were administered. It was hypothesized that the effect of scrambling would be greatest for those programs dealing with topics having the most internal logical development. This hypothesis was not confirmed by the data, however. The investigators cited a number of possible sources of bias that were not accounted for by the study: *redundancy* (i.e., the extent to which a frame consists of sufficient information to stand alone), *repetition* of material in frames so scrambling still is likely to leave at least one frame with cues to the meaning of others, etc.

Pyatte (1969) argued that the lack of more definitive information about the effect of sequence changes on variables such as achievement, retention, and transfer could be attributed in part to a neglect of clear specification of what an ordered sequence of materials is to be; this lack makes it impossible to decide whether a sequence purported to be ordered does meet this condition, and whether a scrambled version of the sequence fails to meet it. In an attempt to follow up this idea, Pyatte (1969) conducted a study in which he defined an ordered sequence as *structured* or *hierarchical*. Assuming that in the hierarchy each point provides positive transfer to the next point, he considered the extent to which positive transfer was acting within a program as a measure of the extent to which the program was hierarchical, and, hence, ordered. His study was designed to provide a check on the effectiveness of the instructional materials, to provide a check on the ordered (structured) materials by examining the differences between these and the unordered (unstructured) materials, and to test the hypothesis that no differences in achievement or transfer would be found between students taking the structured materials and



those taking the unstructured materials. Pyatte reported that both versions of the instructional unit were judged effective; the structured unit was judged to have the defined structure, and the unstructured unit was judged not to have it; and no significant differences were found between the means on achievement or transfer for students taking the structured unit and those taking the unstructured unit. Among the statistically significant findings of the study was evidence that students of high basic ability reach higher levels of achievement and transfer knowledge than students of low basic ability, regardless of mode of program. In concluding his analysis of the study, Pyatte (1969, p. 260) stated that "the effects of sequence on learning measures should at this time be abandoned in favor of attempts to write programs which conform to a defined pattern and to develop the appropriate tools for testing these programs. . . . Having batteries of such well-defined programs, one then would be equipped with the requisite tools for answering questions about the type of program and its effects on such measures as achievement, retention, and transfer."

### *Structure and Traversing Rules*

An interesting early attempt to develop a system for the programmer to follow in structuring sequences was the Ruleg system described by Evans, Homme and Glaser (1962). Glaser (1962, p. 70) stated that this system is "one rough attempt to analyze a knowledge domain, prior to the development of an instructional program, to facilitate the preparation of such a program." The Ruleg technique involves the preparation of a matrix of "rules" which is intended to yield all possible permutations of rules and principles to be included in a program sequence. In a crude way, such a matrix provides information about content structure and possibilities for ordering it. The matrix also could serve as a good source of ideas for formulating instructional hypotheses and documentation procedures. In a study designed to test the feasibility of using the Ruleg system to construct a program by "formula," Evans, Glaser and Homme (1962, p. 451) reported that "Systematically constructed programs can produce, in less learning time, criterion performance comparable with that of a less systematic program." The instructional programs used in the foregoing study were designed to teach the construction of short deductive proofs in symbolic logic.

In a study reported by Newton and Hickey (1965), sequence was specified with respect to a tree-graph structuring of concepts and sub-concepts. Three variables were investigated: *order* or placement of a sub-concept relative to other subconcepts to be learned; *position*—learning all subconcepts first or learning them as they were needed; and *directionality*—whether the statement of the principle or rule which defines a concept is presented before or after the meaning (or generality) of the principle

being developed. In summarizing the results, the authors reported that the foregoing sequence variables did not influence the number of errors made during learning; that performance was faster when principles were stated first and when subconcepts were learned together rather than separately; and that the subjects made significantly poorer scores on a multiple-choice criterion test under conditions in which the subconcepts were most remote from the point where they were integrated to form the major concept. Hickey and Newton (1964) also constructed a potentially fruitful framework for the study of structure and sequence; they set forth a set of eight hypotheses that are worthy of examination and further development.

Scandura (1966a, 1966b, 1966c) conducted a set of three experiments on effects of structurally-related classes of prior learning on familiar and transfer problem solving. Independent variables were: content structure, prerequisites (components of the content structure) available to the subjects, practice on prerequisite components, practice on criterion problem-solving tasks, presentation order, and amount of prelearned prerequisite material. The learning material involved card tasks similar to those common in studies of concept formulation. Significant effects were attributed to presentation order, prerequisite information, and prerequisite practice on problem-solving performance under certain circumstances. Of particular interest was Scandura's effort to carefully construct and explicate the content structure and to take that structure into account in examining outcomes.

Reynolds et al. (1964) studied the variables of repetition and spaced review, and in a statement of rationale for the study asserted that "An understanding of the effect of these variables upon learning from a programmed sequence would have important consequences for the construction of programmed materials, particularly if a set of rules could be established regarding the optimal use of each variable" (p. 1). The findings included: spaced review produced significant increases in learning which persisted, and even increased over a three-week retention interval, and repetition did not produce increased learning or retention.

Short and Haughey (1967) conducted a study to evaluate two contrasting sequencing strategies for teaching multiple-discrimination tasks and to determine if these strategies give consistently different results when they are applied to different subject matter presented in different media. The two sequencing strategies were: *multiple concept*, which consisted of presenting simple descriptions of several related concepts at the beginning and then gradually introducing increasingly complex material pertaining to all of these concepts, and *single concept*, which consisted of presenting one concept at a time and proceeding from a simple description of the concept to a more complex description of it. Only after the concept has

been presented in all its detail and complexity is another concept introduced. The authors reported that in using science materials, students who received the multiple concept sequence performed significantly better than those who received the single concept sequence; no significant differences were found in the comparisons of the foregoing teaching strategies when the students were using language arts materials.

### *Conclusions and Recommendations*

It seems reasonable to conclude that the extent of substantive knowledge about construction of efficient instructional sequences in mathematics is at present desperately sparse. Not nearly enough is known about the connection between the logical structure of the knowledge and the psychological processes involved in acquiring the knowledge. Adequate teaching algorithms which specify the steps to be taken in order to construct an instructional sequence in the presence of a given set of educational ends and a given set of circumstances, and with some assurance of efficiency, do not exist. The systematic and disciplined programs of research and development necessary to construct, evaluate and refine potentially fruitful algorithms have not yet been developed.

To engage in fruitful research on sequence, certain conditions must be met. One of these is the capability of replication of investigations over different occasions of instruction and over different content. There must also be developed long-term purposeful and coherent *strategies* for sequence theory construction and evaluation. Unfortunately, a careful analysis of the research on sequencing provides little evidence that either of these requirements has been adequately dealt with.

### *Bibliography*

- Ausubel, David P. *The Psychology of Meaningful Verbal Learning*. New York: Grune and Stratton, 1963.
- Ausubel, David P.; Robbins, Lillian C.; and Blake, Elias, Jr. Retroactive Inhibition and Facilitation in the Learning of School Materials. *Journal of Educational Psychology* 48: 334-43; 1957.
- Bolvin, John O. Implications of the Individualization of Instruction for Curriculum and Instructional Design. *Audiovisual Instruction* 13: 238-42; 1968.
- Briggs, Leslie J. *Sequencing of Instruction in Relation to Hierarchies of Competence*. Pittsburgh: American Institutes for Research, 1968.
- Evans, James L.; Glaser, Robert; and Homme, Lloyd E. An Investigation of "Teaching Machine" Variables Using Learning Programs in Symbolic Logic. *Journal of Educational Research* 55: 433-52; 1962.
- Evans, James L.; Homme, Lloyd E.; and Glaser, Robert. The Ruleg System for the Construction of Programmed Learning Sequences. *Journal of Educational Research* 55: 513-18; 1962.
- Gagné, Robert M. The Acquisition of Knowledge. *Psychological Review* 69: 355-65; 1962.
- Gagné, Robert M. Learning and Proficiency in Mathematics. *The Mathematics Teacher* 56: 620-26; 1963.

- Gagné, Robert M. *The Conditions of Learning*. New York: Holt, Rinehart, and Winston, 1965.
- Gagné, Robert M. Curriculum Research and the Promotion of Learning. *Perspectives of Curriculum Evaluation*. (Edited by Robert E. Stake.) AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally, 1967a. Pp. 19-38.
- Gagné, Robert M. Training and Principles of Learning. *Studies in Personnel and Industrial Psychology*. (Edited by E. A. Fleishman.) Homewood, Ill.: The Dorsey Press, 1967b.
- Gagné, Robert M. Learning Hierarchies. *Educational Psychologist* 6, No. 1: 3-6; 1968.
- Gagné, Robert M. and Paradise, N. E. *Abilities and Learning Sets in Knowledge Acquisition*. Psychological Monographs: General and Applied, Vol. 75, No. 14 (Whole No. 518). Washington, D.C.: American Psychological Assoc., 1961.
- Gagné, Robert M. et al. *Factors in Acquiring Knowledge of a Mathematical Task*. Psychological Monographs: General and Applied, Vol. 76, No. 7 (Whole No. 526). Washington, D.C.: American Psychological Assoc., 1962.
- Glaser, Robert. Some Research Problems in Automated Instruction: Instructional Programming and Subject-matter Structure. *Programmed Learning and Computer-based Instruction*. (Edited by J. E. Coulson.) New York: Wiley, 1962. Pp. 67-85.
- Hickey, Albert E. and Newton, John M. *The Logical Basis of Teaching: I. The Effect of Subconcept Sequence on Learning*. Newburyport, Mass.: ENTELEK Inc., 1964.
- Hull, T. E. *Introduction to Computing*. Englewood Cliffs, N. J.: Prentice Hall, 1966.
- Kersh, Bert Y. Engineering Instructional Sequences for the Mathematics Classroom. *Research in Mathematics Education*. (Edited by J. M. Scandura.) Washington, D.C.: National Council of Teachers of Mathematics, 1967.
- Merrill, M. David. Correction and Review on Successive Parts in Learning a Hierarchical Task. *Journal of Educational Psychology* 56: 225-34; 1965.
- Merrill, M. David and Stolurow, Lawrence M. Hierarchical Preview Versus Problem Oriented Review in Learning an Imaginary Science. *American Educational Research Journal* 3: 251-61; 1966.
- Newton, John M. and Hickey, Albert E. Sequence Effects in Programmed Learning of a Verbal Concept. *Journal of Educational Psychology* 56: 140-47; 1965.
- Payne, David A.; Krathwohl, David R.; and Gordon, John. The Effect of Sequence on Programmed Instruction. *American Educational Research Journal* 4: 125-32; 1967.
- Popper, Karl R. *Logic of Scientific Discovery*. New York: Basic Books, 1961.
- Pyatte, Jeff A. Some Effects of Unit Structure of Achievement and Transfer. *American Educational Research Journal* 6: 241-61; 1969.
- Reynolds, James H. et al. *Repetition and Spaced Review in Programmed Instruction*. Wright-Patterson Air Force Base, Ohio: Behavioral Sciences Lab., Dec. 1964.
- Roe, Arnold. Comparison of Branching Methods for Programmed Learning. *Journal of Educational Research* 55: 407-16; 1962.
- Roe, K. V.; Case, H. W.; and Roe, Arnold. Scrambled Versus Ordered Sequence in Autoinstructional Programs. *Journal of Educational Psychology* 53: 101-104; 1962.
- Russel, Bertrand. *Human Knowledge: Its Scope and Limits*. New York: Simon and Schuster, 1962.
- Scandura, Joseph M. Algorithm Learning and Problem Solving. *Journal of Experimental Education* 34, No. 4: 1-6; 1966a.
- Scandura, Joseph M. Prior Learning, Presentation Order, and Prerequisite Practice in Problem Solving. *Journal of Experimental Education* 34, No. 4: 12-18; 1966b.
- Scandura, Joseph M. Problem Solving and Prior Learning. *Journal of Experimental Education* 34, No. 4: 7-11; 1966c.
- Scandura, Joseph M. and Wells, Jay N. Advance Organizers in Learning Abstract Mathematics. *American Educational Research Journal* 4: 295-301; 1967.
- Schwab, Joseph J. *The Concept of Structure in the Subject Fields*. Paper presented at the Twentieth Annual Meeting of the Council on Cooperation in Teacher Education of the American Council on Education, October, 1961, Washington, D.C. Chicago: Univ. of Chicago. (Mimeo.)
- Short, J. and Haughey, B. Study of Sequencing Strategies. *Proceedings of the 75th*



- Annual Convention of the American Psychological Association*, Vol. 2. Washington, D. C.: American Psychological Assoc., 1967. Pp. 317, 318.
- Suppes, Patrick. Mathematical Concept Formation in Children. *American Psychologist* 21: 139-50; 1966.
- Suppes, Patrick. Some Theoretical Models for Mathematics Learning. *Journal of Research and Development in Education* 1: 5-22; 1967.
- Suppes, Patrick and Groen, Guy. Some Counting Models for First-Grade Performance Data on Simple Addition Facts. Psychology Series, Institute for Mathematical Studies in the Social Sciences, Technical Report No. 90. Stanford, Calif.: Stanford Univ., 1966.
- Suppes, Patrick; Hyman, Lester; and Jerman, Max. Linear Structural Models for Response and Latency Performance in Arithmetic on Computer-Controlled Terminals. Psychology Series, Institute for Mathematical Studies in the Social Sciences, Technical Report No. 90. Stanford: Stanford Univ., 1966.
- Suppes, Patrick; Jerman, Max; and Dow, Brian. Computer-Assisted Instruction: Stanford's 1965-66 Arithmetic Program. New York: Academic Press, 1968.
- Woodward, E. L. Comparative Study of Teaching Strategies Involving Advance Organizers and Post Organizers and Discovery and Nondiscovery Techniques Where the Instruction is Mediated by Computer. Doctor's thesis. Tallahassee: Fla. State Univ., 1966.

## AUTHOR

RALPH T. HEIMER Address: Pennsylvania State University, University Park, Pennsylvania Title: Associate Professor of Education Age: 40 Degrees: B.S., Lockhaven State College; M.A. and Ed.D., Pennsylvania State Univ. Specialization: Instructional Theory Development With Special Reference To Sequencing.

## 9: ACTIVITY LEARNING

Thomas E. Kieren\*

*University of Alberta*

This review of activity learning in mathematics covers publications from January 1964 through December 1968. In the context of this article *activity learning* is taken to mean school learning settings in which the learner develops mathematical concepts through active participation. This process may involve the manipulation of physical materials, the use of games, or partaking in experiments with physical objects. Theory and research on the effects of this type of activity will be discussed in this article under manipulative learning.

The second form of activity learning discussed here is presented under the rubric of discovery learning. In this form of learning, the student actively engages in the process of forming mathematical ideas for himself, but does not necessarily use any physical devices in learning mathematics. This review is generally restricted to those studies concerned with classroom instruction in mathematics; it does not include research conducted basically for psychological considerations.

There were quite a number of empirical studies, particularly in discovery learning, in which comparatively standard research methodology was used. However, much of the research in manipulative learning in mathematics would fit into Johnson's (1966) categories of action-naturalistic or philosophical research.

The discovery research reviewed was centered on the North American continent. However, development and research work in manipulative material and games-oriented learning in mathematics seem to be worldwide in scope (Weaver, 1967). Work has been completed in England (E. Biggs, 1965, 1968), France (Glaymann, 1967), Australia (J. Biggs, 1965; Golding, 1968), and in the Canary Islands (Capparros and Delgado, 1967) as well as in the United States and Canada.

### *Discovery Teaching and Learning in Mathematics*

There was no lack of published written work defending or attacking the effectiveness of discovery methods of instruction. As Cronbach (1966)

---

\*Dr. D. B. Harrison, University of Calgary, served as consultant to Dr. Kieren on the preparation of this chapter.

pointed out there are a multitude of different approaches that have been applied under the rubric of "discovery." Thus, the theoretical discussion and the research in discovery learning suffered from lack of precision and lack of communication caused by the variations in connotations of "discovery." As part of a conference report (edited by Shulman and Keislar, 1966), Davis (1966) attributed several positive effects to a discovery learning experience, especially in a classroom group. He suggested that the child is rewarded for his efforts by the pleasure of his own discovery and his ability to bring closure to a problem setting himself, that the child is motivated by the opportunity to display his ideas in competition with others, and that the student learns to act independently in checking the veracity of his ideas. Bruner (1966a) concluded that discovery settings present opportunities for students to draw for themselves relationships between things they know and the learning task at hand. In addition, discovery methods build problem-solving skill by providing experience in pushing ideas to their logical limit and in effectively forming concise hypotheses. Bruner, like Davis, suggested that thinking acts are reinforced by the discovery accomplished and that a reflective attitude is developed in students.

Ausubel, in a counter-argument to those of Bruner and Davis, stressed that discovery processes are less efficient than didactic or expository processes and that discovery is not needed for meaningful acquisition of knowledge. Bittinger (1968) cited literature suggesting that discovery attitudes can be sustained via didactic teaching. He further noted Cronbach's hypothesis that discovery methods may not work well for anxious, dependent students. Becker and MacLeod (1967), in reviewing literature relating discovery and transfer, found no conclusive research evidence that discovery methods foster transfer. This finding suggests that discovery methods might not facilitate the acquisition of superior problem-solving capabilities. Nonetheless, Becker and MacLeod found that the literature indicates a conflict between methodologies that sponsor maximum understanding and those that sponsor maximum motivation to continue learning.

The research published during 1964-1968 on discovery methodologies in mathematics instruction did little to resolve the theoretical conflict outlined above. The results presented a mixed picture, and the discovery methodology used was frequently not clearly defined. When the definition was careful, discovery and its experimental opposite were most frequently defined in terms of sequence variables alone. In research conducted with 432 fifth and sixth grade children, Worthen (1967, 1968) carefully defined discovery and expository modes of presentation. The discovery method used an examples-rule sequence; the expository teaching always began with a verbalization of the rule. Controls were exercised over the teacher variable, and checks were made to ascertain that different methods were

maintained by the teachers in their two classes. Materials used were equivalent in length and numbers of examples and exercises. Worthen found that the Expository group was significantly better on a Concept Knowledge Test, but the Discovery Group was significantly better, at least at the .10 level of significance, on short and long-term retention and transfer. Eldredge (1965), in a related study using programmed materials with the same sequencing differential between groups, found that the discovery group whose method was defined as a hints-then-formula method did significantly better than the expository group (rule-examples) on learning how to sum number series. Roughead and Scandura (1968) found that "rule only" and "rule then discovery of why rule works" strategies were significantly worse than strategies using other sequences, such as examples-rule. The authors suggested that learning how an algorithm works may interfere with the later learning of why the algorithm works. Meconi (1967a, 1967b) worked with a sample of 45 high-ability ninth grade students; he used three methods: rule-example, guided discovery using leading examples, and pure discovery. Meconi found no differences in learning or retaining problem-solving ability; however, the pure discovery group was significantly more efficient at learning to solve new problems. Guthrie (1967), with a sample of 72 college seniors, found that a rule-example group was significantly better in learning to solve cryptograms than an example-rule, an example-only, or a control group. On rule transfer, the example-only group was superior to the example-rule and control groups, which were in turn superior to the rule-example group. In summary, there appears to be an advantage in an example-rule sequencing in the transfer of learning, but with initial performance and retention the method appears to interact with maturity of the learners.

Bittinger (1968) cited studies in which he found superiority for discovery groups on discovery tests; he also found that students learning via inquiry methods were at least as good in achievement as students learning via an expository approach and were significantly more fluent in generating potential solutions to problems. Glennon and Callahan (1968) cited a study by Carlow on the effects of consolidation of discovered learning. Carlow used a random sample of 36 college-preparatory ninth graders. Each student was taught individually and given a sequence of hints as he needed them while learning generalizations in probability. Half the group was given consolidation work with exercises on learned materials. Since initial learning was to a criterion, Carlow employed a retention-transfer test on generalizations in probability. On this test there was a significant mean difference favoring the group given the consolidation work. This difference occurred in groups using two different mathematical strategies, ordered partitions and permutations-combinations. In addition, certain personality factors, including submissiveness, and con-



ceptual level seemed to relate to effectiveness of discovery methods. Retzer and Henderson (1967) used a sample of eighty junior high students with IQs above 116. One group received training in principles of logic but the second group did not. A design using a two-way classification (logic vs. non-logic and gifted vs. non-gifted) showed that in verbalizing discovered principles, the students trained in logic did significantly better than those not trained in logic. There was a significant interaction of the treatment variable and intelligence which suggested that the gifted were more aided by the study of logic than were the non-gifted.

Whether discovery methods promote creativity in mathematics has been looked at in a much less formal way. In citing case histories of two bright children, Biggs (1968a) and Davis (1967b) showed instances in which discovery settings fostered vastly creative enterprises of grade school children. Biggs cited the case of Peter, who discovered principles of differential calculus from a study of graphs of data from experiments relating angle of inclination of a track to the speed of a toy car. Davis discussed the discovery of an original subtraction algorithm involving negative numbers by a third grader named Kye. Davis ascribed the discovery in part to the open atmosphere provided by the discovery-oriented setting in which the teacher listened to student responses without prejudging them.

Two approaches have been researched on discovery and its effects on low achievers' achievement in and attitudes toward mathematics. Berger and Howitz (1967) chose a nationwide sample of 29 teachers, each with two classes of general mathematics students in grade 9. One class used *Experiences in Mathematical Discovery*, written materials designed to stimulate a discovery approach to the learning of modern mathematics concepts by low achievers. The control classes used a variety of remedial and consumer arithmetic texts. The SCAT(3A) and the STEP(3A) tests were administered as pretests. Because of the marked variability in class performances, 6 pairs of classes were analyzed separately and the remaining 23 pairs classified by the mean SCAT(3A) score into five groups on which separate analyses were done. The investigators then classified subjects according to treatment, performance on STEP (3A) test, and teacher. Using analyses said to be appropriate to this multifactor arrangement of treatments, the investigators found no significant treatment effects as measured by the STEP(3B) test or an attitude measure. As might be expected, there was a significant effect favoring the experimental group on a test covering *Experiences in Mathematical Discovery*. This treatment effect, however, was confounded by a significant treatment by teacher interaction. Price (1967) used three classes of tenth grade students to examine the effects of two discovery methods on low achievers in mathematics. One class was guided to make discoveries of generalizations while

a second class received the same instruction but was guided in applying generalizations to practical problems. The third class served as a control group in which more traditional patterns of instruction were used. All classes had the same teacher. All classes showed gains in algebra achievement. However, the discovery and transfer classes were significantly better than the control class in pretest to posttest gains on The Survey Test of Algebra Aptitude. The transfer class showed significantly greater pretest to posttest gain than did the other two classes on the Watson-Glaser Critical Thinking Appraisal. Price concluded that the discovery approach had in itself no significant transfer-generating capacity. In attitude measured by student ranking of mathematics with other subjects, the two treatment groups showed a positive attitude change while the control groups exhibited a negative attitude change.

*Additional References:* Ballew (1965); Caruso (1966); Crabtree (1966); Craig (1965); Davis (1967); Fleckman (1966); Friedlander (1965); Hanson (1967); Henderson and Rollins (1967); Howitz (1965); Levine (1967); Scandura (1964, 1968).

### *Manipulative Learning in Mathematics*

During 1964-1968 there was no dearth of theoretical discussion on the value of manipulative learning in mathematics. Bruner (1966b) suggested that the optimal sequence for learning mathematical ideas includes enactive, iconic or image manipulative, and symbolic stages. Although secondary school children can operate at the symbolic level alone, Bruner sees a danger in simply instructing and learning at this level: if in learning or problem solving symbolic transformation fails a person, he needs to be able to function with appropriate manipulative or iconic transformations. In his discussion of the application of the work of Piaget, Isaacs (1964) cited the use of the environment as a natural stimulus for mathematical ideas. He further suggested the value of having a great variety of equipment and materials in the learning environment and the need for specially enriching environments for children coming from sterile environments. Davis (1967a) observed that learning from physical materials added reality to the learning situation and provided an alternative to authoritarian teaching.

Dienes (1967) suggested fostering concept development through the use of multiple concrete and game embodiments of a concept. He noted the need for perceptual and mathematical variates of a concept in its development for children. E. E. Biggs discussed the research background for activity learning in mathematics. She cited Hebb on the need for a stimulating environment, Lunzer on the role of intuitive reasoning in the metamorphoses of reasoning, and Tough on the significance of con-

crete experiences and related language development in the development of number concepts in preschool children. J. Biggs (1965) claimed superiority for a multimodel environment for mathematics learning over a unimodel environment, which uses only one type of manipulative materials such as Cuisenaire rods.

The quality of the actual research and its attention to the questions raised above are questionable. Many of the studies were pilot studies in some sense, and questions were often asked that were not sufficiently complex to deal with the theoretical and practical issues. Careful definition of the problem was frequently lacking. Nevertheless, the research cited in the following pages represents an effort to establish the contribution of more playlike approaches to mathematics learning.

Sutton-Smith (1967) discussed research on the role of play in cognitive development. He found research that supported play as a significant sponsor of novel responses to creativity test settings and as a mode for increasing the child's response repertoire. From Piaget's work, Sutton-Smith concluded that play allows for responses in fantasy that cannot be made in reality. This conclusion was supported by Dienes (1964), who suggested the importance of the role of elemental play and mathematical rule-based games in mathematics learning in the elementary school. Via such games the student may cope with advanced mathematical ideas in an appropriate manner. Sutton-Smith also found that learning via play correlates highly with information-seeking behavior.

Williams (1967) and J. Biggs (1965) separately reported on a three-year study in England and Wales in which three structural methods were compared with traditional methods of teaching in the early elementary school; the three structural methods were: the Cuisenaire and the Stern, which are unimodel methods, and the Dienes, which is a multimodel method. The experimenters either compared students in separate schools or compared students using a structural method to previous students in the school who used traditional methods. Williams documented the statistical controls for student ability and the efforts in controlling for teacher experience, constancy of method, and the Hawthorne effect in this longitudinal study. Biggs found that high IQ boys fared better on understanding, motivation and attitude under unimodel methods than under traditional methods. The traditional methods appeared to produce good computational facility but high number anxiety. The multimodel group was superior to either of the other two groups on the three criteria mentioned above.

Most other research dealing with the use of manipulative materials compared groups using the Cuisenaire method to control groups using more traditional methods. Robinson (1964) and Nelson (1964) discussed fifty comparisons involving 20,000 students in grades 1 to 6 in Canada.

They found that the results showed that symbol manipulation facility was far in advance of comprehension of practical significance of the computation. They found evidence which suggested that the Cuisenaire groups were significantly poorer on word problems and that these groups had less ability in transferring their knowledge of mathematics to more generalized arithmetic settings.

There were several other studies—Luow (1964), Hollis (1965), and Nasca (1966)—that supported the contention that the Cuisenaire method group has advanced computational ability in early elementary school.

The positive results cited above conflict with those of Passy and Brownell reported in an earlier review by Weaver and Gibb (1964). The results should be considered in the light of Nelson's (1964) criticisms applied to Canadian research: samples were unrepresentative; the results might be an artifact of the Hawthorne effect or use of gain scores; and conclusions go beyond the data.

The role of games in learning mathematics was discussed in three studies. Anderson (1965) had a random sample of first grade students play programmed games with experimenters for 27 sessions. These games were based on logical skills. The experimental group was compared with a randomly selected control group on the criterion test of novel problems involving conjunctive logical statements. The experimental group was found to be statistically superior to the control group in number of problems solved and transfer problems solved and needed fewer trials to solve problems on the test. Humphrey (1965) reported an exploratory study which suggested that first graders using active games exhibit greater gains in the learning of number concepts than students using a workbook to study the same concepts.

Davis (1967b), in his comprehensive report of the Madison Project, stated that the goals of the project were to prepare classroom experiences developed in a specific manner, to train teachers in the use of the experiences, and to test the lessons. These experiences were, among other things, designed to provide a foundation for later work in mathematics, emphasize mathematics as a process, provide greater use of physical materials, and develop an atmosphere receptive to student initiative. Hopefully, these experiences would then provide understanding of mathematical concepts and unify arithmetical concepts while arousing interest and relating mathematics to science. The goal of the project was to produce curricula of viable classroom experiences. The methodology of evaluation of experience called for trial and criticism, revision, and then retrieval with eventual recording via video or audio tapes. These tapes were used to help train teachers so the viability of the procedures could be tested on a broad population. Given this frame of reference, results of the project are diffi-



cult to report in statistical terms. Davis stated that five curricula have been produced. The curriculum designed for bright junior high students is the only one that has proven unstable over a large number of classes. The later curricula have changed their emphases from large group discussions to small group settings. Cleary (1967) used 77 students in three classes, two from schools using the Madison Materials Project and one from a similar school which did not. The Madison Project students were in seventh grade; the control group consisted of seventh grade students who had studied in a setting using modern mathematics materials. The investigator chose 45 items relevant to Madison Project lessons from a 400-item pool for the Stanford Achievement High School Test Battery. The mean scores of the two Madison Project classes were higher than the control class mean. However, a *t* statistic showed only one Madison Project class mean to be significantly higher than the control class mean. Interviews with a sample of 15 students indicated that the control group superiority on certain items on quadratic equations apparently occurred because students were able to obtain correct solutions simply through substitution in expressions. Davis (1967b) reported an interview study by a clinical psychologist, H. Barrett, done in connection with the Madison Project. After a year of conducting interviews, Barrett concluded that students in grades 6 and 7 like physical activity and courses involving it and do not like doing sophisticated mental tasks. Barrett's data formed the basis for the Madison Project decision to use physical materials and laboratories at grades K-9 and also in college courses for prospective mathematics teachers.

*Additional References:* Callahan and Jacobson (1967); Davis (1966); Fitzgerald (1968); Golding (1968); Gowder (1965); Harrison (1968); Harvin (1965); Hopkins (1965); Jones (1968); Kaplan (1964); Kieren and Vance (1968); Moise (1966); Parkyani (1967).

### *Critique of the Research Reviewed*

As I suggested at the outset of this review, the literature on activity learning during the 1964-1968 period could be characterized as theoretical discussions or descriptions of developmental efforts. Many of the studies mentioned above were only pilot projects for followup studies hopefully in progress today. Hence, most of the studies were small in scale, and perhaps far too lacking in control and in potential generalizability to be considered good research. Nevertheless, they represent first steps toward answering the complex question, "what is the effect of activity methodologies on the learning of mathematics?" The studies in discovery and manipulative methodology dealt with persons from early school age through college age. Studies focused on a range of student abilities, although inclusion of the range of abilities did not occur on a systematic basis in any

of the research reviewed. The research reviewed also showed efforts to develop underlying theoretical models for learning by discovery, for example, the work of Scandura, and for manipulative learning, as illustrated by the efforts of J. Biggs (1965) and Dienes and Jeeves (1965). The massive work of Davis and that of the Nuffield Project (Matthews, 1965; E. Biggs, 1968b) represented attempts to immediately influence inschool practices through the development and testing of operational ideas for activity learning.

But for all of these encouraging aspects and rather favorable results, the research reviewed can generally be criticised in the manner of Cronbach (1966). The discovery and manipulative activity research generally suffered from the bias of having the development effort expanded on the discovery or manipulative treatment. Further, the research cited for the most part asked the simple question, "Is activity learning better than something else?" Cronbach (1966) cited the need instead to study a five-fold interaction among subject matter, instructional type, timing, type of pupil, and desired outcomes. Most of the studies are also liable to Wittrock's (1966) criticism. He suggested that most discovery studies conducted prior to 1966 lacked replicability, dealt with intact classes, and lacked operational definitions of the settings and treatments. He further suggested that discovery needs to be studied as an independent *and* as a dependent variable and that these two uses should not be confused. All of Wittrock's suggestions and criticism certainly can apply to studies in manipulative methods. Regarding Wittrock's last suggestion, it is interesting to note that, with the exception of the study by Dienes and Jeeves (1965), none of the manipulative studies used ability in activity settings as a criterion measure or as a dependent variable.

Davis (1967a) questioned the usefulness of the whole formal rhetoric of coded generalizations at the current stage of research in activity learning. He suggested that there is a need for research to develop practitioner's maxims and used a methodology, discussed in his report of the Madison Project (Davis, 1967b), which allowed him to study interaction of several of the parameters suggested by Cronbach. However, Davis can be criticized for not carefully defining some of his terms, such as *viability*. Thus, his results are hard to judge and one is left with the feeling that they might be unstable across several observers studying the designed experiences.

*Additional References:* Brownell (1966); Sawada and Nelson (1967).

### *Suggestions for Further Research*

The studies in discovery methodology exhibit a trend that favors a sequencing in learning of examples followed by a formal rule. Studies

need to be done to generalize this method to include levels of student activity in defining the problems, developing suitable examples and conjectured solutions, and evaluating conjectures and formalizing the generalizations. In other words, given the effectiveness of a particular sequence, viz., example-rule, a variety of methods of example generation and rule determination should be studied. Obviously, this should be done across several areas of subject matter and several mathematical processes. Researchers should study the relationship of method with several different learning outcomes and learning characteristics. As more knowledge is obtained about cognitive styles, characteristics of students who function well under varying levels of student responsibility in example generation and rule determination should be studied.

Within the context of manipulative methodology, studies need to be done to determine the value of actual as opposed to vicarious manipulation at the upper elementary and junior high school levels. For what purpose and for which students is the laboratory in mathematics an effective adjunct to mathematics classes? What are the contributions of curriculum enriching and curriculum reinforcing laboratory approaches? Can the manipulative experience contribute to various desired learning outcomes in a sequence of experience-generalization-application? How should laboratories be used with low achievers? What are the effects of activity learning experiences provided by the computer? What are the contributions of specialized manipulative experiences to the development of mathematical imagery? For what purposes and for whom is the sequence enactive-iconic-symbolic valuable in learning mathematics? What are the relative contributions of a multimodel setting to mathematics learning and for whom is it most effective? In addition, the work started by Dienes and Jeeves (1965) opens many avenues for replication and further development. Should the learning styles postulated by Dienes and Jeeves prove stable, they may serve as independent or dependent variables in the study of manipulative learning in mathematics.

If researchers are to explore such questions thoroughly enough to get good answers, they will need the cooperative efforts of several other researchers. Thus, even the work of graduate students and school systems needs to be done on some coordinated, cooperative bases. Otherwise, there is little hope of going beyond the first step level of much of the research work reviewed here.

### *Bibliography*

- Anderson, Richard C. Can First Graders Learn an Advanced Problem Skill? *Journal of Educational Psychology* 56: 283-94; 1965.
- Ausubel, D. P. Facilitating Meaningful Verbal Reception Learning in the Classroom. *The Arithmetic Teacher* 15: 126-32; 1968.

- Becker, J. D. and MacLeod, G. K. Teaching, Discovery, and Problems of Transfer of Training. *Research in Mathematics Education*. (Edited by J. Scandura.) Washington, D. C.: National Council of Teachers of Mathematics, 1967. Chapter IX., pp. 93-107.
- Berger, E. J. and Howitz, T. A. Evaluation of Experiences in Mathematical Discovery. *Research in Mathematics Education*. (Edited by Scandura.) Washington D. C.: National Council of Teachers of Mathematics, 1967. Chapter VI., pp. 66-69.
- Biggs, E. E. Research in the Children's Method of Learning. *Mathematics in Primary School. Curriculum Bulletin 1: 5-9*. London: School Council for the Curriculum and Examinations, 1965.
- Biggs, E. E. Mathematics Laboratories and Teacher Centres—The Mathematics Revolution in Britain. *The Arithmetic Teacher* 15: 400-408; 1968a.
- Biggs, E. E. The Teaching of Mathematics in the Primary School. *Reports of the Commonwealth Conference of Mathematics Education*. London: Commonwealth Secretariat, 1968b.
- Biggs, J. Towards a Psychology of Educative Learning. *International Review of Education* 11: 77-92; 1965.
- Bittinger, M. L. A Review of Discovery. *The Mathematics Teacher* 61: 140-46; 1968.
- Bruner, J. S. Some Elements of Discovery. *Learning by Discovery: A Critical Appraisal*. (Edited by L. Shulman and E. Keislar.) Chicago: Rand McNally, 1966a. Chapter VII., pp. 101-14.
- Bruner, J. S. *Towards a Theory of Instruction*. Cambridge, Mass: Harvard Univ. Press, 1966b.
- Capparros, J. and Delgado, M. *First Report on the Canary Islands Mathematics Project*. Sherbrooke, Quebec: International Study Group for Mathematics Learning, Centre Psycho-Mathématique, Université de Sherbrooke, 1967.
- Cleary, J. R. A Study of Test Performance in Two Madison Project Schools and One Control School. *A Modern Mathematics Program as it Pertains to the Interrelationship of Mathematical Content, Teaching Methods and Classroom Atmosphere (The Madison Project)*. (Edited by Robert B. Davis.) Cooperative Research Project No. D-233. Syracuse, N. Y.: Syracuse Univ., Oct. 1967. Appendices, A6-A25.
- Cronbach, Lee J. The Logic of Experiments on Discovery. *Learning by Discovery: A Critical Appraisal*. (Edited by L. Shulman and E. Keislar.) Chicago: Rand McNally, 1966. Chapter V., pp. 76-92.
- Davis, Robert B. Discovery in the Teaching of Mathematics. *Learning by Discovery: A Critical Appraisal*. (Edited by L. Shulman and E. Keislar.) Chicago: Rand McNally, 1966. Chapter VIII., p. 115-28.
- Davis, Robert B. The Range of Rhetorics, Scale, and Other Variables. *Journal of Research and Development in Education* 1: 51-68; 1967a.
- Davis, Robert B. (editor). *A Modern Mathematics Program as It Pertains to the Interrelationship of Mathematical Content, Teaching Methods and Classroom Atmosphere (The Madison Project)*. Cooperative Research Project No. D-233. Syracuse, N.Y.: Syracuse Univ., Oct. 1967b.
- Dienes, Z. P. Some Basic Processes Involved in Mathematics Learning. *Research in Mathematics Education*. (Edited by J. Scandura.) Washington, D. C.: National Council of Teachers of Mathematics, 1967. Chapter III., p. 21-34.
- Dienes, Z. P. *The Power of Mathematics*. London: Hutchinson Educational Press, 1964.
- Dienes, Z. P. and Jeeves, M. A. *Thinking in Structures*. London: Hutchinson Educational Press, 1965.
- Eldredge, G. M. Discovery vs. Expository Sequencing in a Programmed Unit on Summing Number Series. *Sequence Characteristics of Textual Materials and Transfer of Learning. Part I, Experiments in Discovery Learning*. (Edited by G. M. Della-Piana; G. M. Eldredge; and B. R. Worthen.) Salt Lake City: Bureau of Education Research, Univ. of Utah, 1965.
- Glaxman, Maurice. *Initiation à la Logique. Rapport pour la Reunion de l'U.N.E.S.C.O.* Sherbrooke, Quebec: International Study Group on Mathematics Learning, Centre Psycho-Mathématique, Université de Sherbrooke, Nov. 1967.
- Glennon, V. J. and Callahan, L. G. *A Guide to Current Research in Elementary School*



- Mathematics. Washington, D. C.: Assoc. for Supervision and Curriculum Development, 1968.
- Golding, E. W. *Report on Teacher Education with Special Reference to Mathematics Learning*. Sherbrooke, Quebec: International Study Group for Mathematics Learning, Centre Psycho-Mathématique, Université de Sherbrooke, Jan. 1968.
- Guthrie, John T. Expository Instruction vs. the Discovery Method *Journal of Educational Psychology* 58: 45-49; 1967.
- Hollis, L. V. A Study to Compare the Effects of Teaching First and Second Grade Mathematics by the Cuisenaire-Gattegno Method with a Traditional Method. *School Science and Mathematics* 65: 683-87; 1965.
- Humphrey, J. H. An Exploratory Study of Active Games in Learning of Number Concepts by First Grade Boys and Girls. *Perceptual and Motor Skills* 23: 341-42; 1965.
- Isaacs, N. Psychology, Piaget, and Practice—Continued. *Teaching Arithmetic* 2: 2-9; Autumn 1964.
- Johnson, D. A. A Pattern for Research in Mathematics Education. *The Mathematics Teacher* 59: 418-28; 1966.
- Lucow, W. H. An Experiment with the Cuisenaire Method in Grade Three. *American Educational Research Journal* 1: 159-67; 1964.
- Matthews, Geoffrey. The Nuffield Mathematics Teaching Project. *The Arithmetic Teacher* 15: 101-102; 1968.
- Meconi, L. J. Concept Learning and Retention in Mathematics. *Journal of Experimental Education* 26: 51-57; 1967a.
- Meconi, L. J. The Mathematically Gifted Student and Discovery Learning. *The Mathematics Teacher* 60: 862-65; 1967b.
- Nasca, Donald. Comparative Merits of a Manipulative Approach to Second Grade Arithmetic. *The Mathematics Teacher* 13: 221-26; 1966.
- Nelson, L. D. Analysis of Canadian Research on the Effectiveness of Cuisenaire Materials. *Canadian Experience with Cuisenaire Method*. Ottawa, Ontario: Canadian Council for Research in Education, the Canada Council (140 Willington Street), 1964. Pp. 183-87.
- Price, Jack. Discovery: Its Effect on Critical Thinking and Achievement in Mathematics. *The Mathematics Teacher* 60: 874-76; 1967.
- Retzer, K. A. and Henderson, K. B. Effect of Teaching Concepts of Logic on Verbalization of Discovered Mathematical Generalizations. *The Mathematics Teacher* 60: 707-10; 1967.
- Robinson, F. G. A note on the Quantity and Quality of Canadian Research on the Cuisenaire Method. *Canadian Experience with the Cuisenaire Method*. Ottawa, Ontario: Canadian Council for Research in Education, the Canada Council (140 Willington St.), 1964. Pp. 181-82.
- Roughead, W. G. and Scandura, J. M. What is Learned in Mathematical Discovery. *Journal of Educational Psychology* 59: 283-89; 1968.
- Shulman, Lee S. and Keislar, Evan R. (editors). *Learning by Discovery: A Critical Appraisal*. Chicago: Rand McNally, 1966.
- Sutton-Smith, B. The Role of Play in Cognitive Development. *The Young Child, Review of Research*. (Edited by W. Hartup and N. Smothergill.) Washington, D. C.: National Assoc. for the Education of Young Children, 1967. Pp. 96-100.
- Weaver, J. F. Action Research and Curriculum Development. *The Arithmetic Teacher* 14: 228-30; 1967.
- Weaver, J. F. and Gibb, E. G. Mathematics in the Elementary School. *Review of Educational Research* 31: 273-83; 1964.
- Williams, J. D. (editor). *The Evaluation of Newly Introduced Teaching Methods. Mathematics Reform in the Primary School*. New York: UNESCO Institute for Education, 1967. Pp. 58-67.
- Wittrock, M. C. The Learning by Discovery Hypothesis. *Learning by Discovery: A Critical Appraisal*. (Edited by L. Shulman and E. Keislar.) Chicago: Rand McNally, 1966. Chapter IV., pp. 33-75.
- Worthen, Blaine R. A Comparison of Discovery and Expository Sequencing in Elementary Mathematics Instruction. *Research in Mathematics Education*. (Edited by J. Scan-

- dura.) Washington, D. C.: National Council of Teachers of Mathematics, 1967. Chapter V., pp. 44-59.
- Worthen, Blaine R. Discovery and Expository Task Presentation in Elementary Mathematics. *Journal of Education Psychology Monograph Supplement* 59, #1, Part 2: Feb. 1968.

### Additional References

- Ballew, Julius H. *A Study of the Effect of Discovery Learning Upon Critical Thinking Abilities of High School Algebra Students*. Doctor's thesis. Chapel Hill: Univ. of N. C., 1965. Abstract: *Dissertation Abstracts* 26: 3775, 1966.
- Brownell, W. A. The Evaluation of Learning Under Dissimilar Systems of Instruction. *The Arithmetic Teacher* 13: 267-74; 1966.
- Callahan, J. J. and Jacobson, R. S. An Experiment with Retarded Children and Cuisenaire Rods. *The Arithmetic Teacher* 14: 10-13; 1967.
- Caruso, George E. *A Comparison of Two Methods of Teaching the Mathematical Theory of Groups, Rings, and Fields to College Freshmen*. Doctor's thesis. New York: N. Y. Univ., 1966. Abstract: *Dissertation Abstracts* 27: 3769, 1967.
- Crabtree, C. The Effects of Structuring on the Productiveness of Children's Thinking. *Journal of Experimental Education* 35: 1-13; 1966.
- Craig, Robert C. Discovery, Task Completion, and the Assignment as Factors in Motivation. *American Education Research Journal* 2: 217-20; 1965.
- Davis, Robert B. The Next Few Years. *The Arithmetic Teacher* 13: 355-62; 1966.
- Davis, Robert B. *The Changing Curriculum: Mathematics*. Washington, D. C.: Assoc. for Supervision and Curriculum Development, 1967.
- Fitzgerald, W. M. A Mathematics Laboratory for Elementary School Teachers. *The Arithmetic Teacher* 15: 547-49; 1968.
- Fleckman, Bessie. *Improvement of Learning Division Through Use of the Discovery Method*. Doctor's thesis. Athens: Univ. of Ga., 1966. Abstract: *Dissertation Abstracts* 26: 4329, 1966.
- Friedlander, B. F. A Psychologist's Second Thoughts on Concepts, Curiosity and Discovery in Teaching and Learning. *Harvard Educational Review* 35: 18-38; 1965.
- Gowder, Alex B. *A Comparative Study of Two Methods of Teaching Arithmetic in the First Grade*. Doctor's thesis. Denton: N. Tex. State Univ., 1965. *Dissertation Abstracts* 26: 3778; No. 7, 1966.
- Hanson, L. E. *Inductive Discovery Learning, Reception Learning and Formal Verbalization of Mathematical Concepts*. Doctor's thesis. Tallahassee: Fla. State Univ., 1967. Abstract: *Dissertation Abstracts* 28: 1731A-32A; No. 5, 1967.
- Harrison, D. B. Piaget, Bruner and the Teacher. *Manitoba Journal of Education* 4: 7-19; Nov. 1968.
- Harvin, Virginia R. Analysis of the Uses of Instructional Materials by a Selected Group of Teachers of Elementary School Mathematics. *The Arithmetic Teacher* 12: 562; 1965.
- Henderson, K. B. and Rollins, J. H. A Comparison of Three Strategems for Teaching Mathematical Concepts and Generalizations by Guided Discovery. *The Arithmetic Teacher* 19: 583-88; 1967.
- Hopkins, C. D. *An Experiment on Use of Arithmetic Time in the Fifth Grade*. Doctor's thesis. Bloomington: Univ. of Ind., 1965. Abstract: *Dissertation Abstracts* 26: 5291, 1966.
- Howitz, Thomas A. *The Discovery Approach: A Study of its Relative Effectiveness in Mathematics*. Doctor's thesis. Minneapolis: Univ. of Minn., 1965. Abstract: *Dissertation Abstracts* 26: 7178; No. 12, 1966.
- Jones, T. The Effect of Modified Programmed Lectures and Mathematical Games Upon Achievement and Attitude of Ninth Grade Low Achievers in Mathematics. *The Mathematics Teacher* 61: 603-607; 1968.
- Kaplan, J. D. An Example of Student Generated Sequences in Mathematics Instruction. *The Mathematics Teacher* 57: 298-302; 1964.

- Kieren, T. E. and Vance, J. H. The Theory of Active Learning: Its Application in a Mathematics Workshop. *Manitoba Journal of Education* 4: 33-40; Nov. 1968.
- Levine, Joan L. A Comparative Study of Two Methods of Teaching Mathematical Analysis at the College Level. Doctor's thesis. New York: Teachers College, Columbia Univ., 1967. Abstract: *Dissertation Abstracts* 28: 4052, 1968.
- Moise, Edwin E. Activity and Motivation in Mathematics. *Mathematics Teaching* 34: 64-67, Spring 1966.
- Parkanyi, B. On the Construction of an Apparatus for Learning Some Group Structures. Sherbrooke, Quebec: International Study Group for Mathematics Learning, Centre Psycho-Mathématique, Université de Sherbrooke, 1967.
- Sawada, Daiyo and Nelson, L. D. Conservation of Length and the Teaching of Linear Measurement: A Methodological Critique. *The Arithmetic Teacher* 14: 345-48; 1967.
- Scandura, J. M. An Analysis of Exposition and Discovery Modes of Problem Solving Instruction. *Journal of Experimental Education* 33: 149-60; 1964.
- Scandura, J. M. Research in Psychomathematics. *The Mathematics Teacher* 61: 581-91; 1968.

## AUTHOR

- THOMAS E. KIEREN Address: Pennsylvania State University, University Park, Pennsylvania Title: Assistant Professor of Education Age: 28 Degrees: B.A., Carleton College; M.A., Northwestern Univ.; Ph.D., Univ. of Minnesota Specialization: Investigating the Effects of Components of Multi-media Settings on Mathematics Learning with Special Reference to the Uses of Computers and Mathematics Laboratories.

## 10: PROBLEM SOLVING IN MATHEMATICS

Jeremy Kilpatrick\*

*Teachers College, Columbia University*

The preeminence of increased problem-solving ability as a goal of mathematics instruction has long been admitted; but like the weather, problem solving has been more talked about than predicted, controlled, or understood. The studies reviewed in this chapter were chosen from a large number published during the last five years that are relevant to the twin issues of how problem solving is learned and how it can be taught. The role of problems in developing students' mathematical activity was chosen by the International Commission on Mathematical Instruction as one of three topics for discussion at the 1966 International Congress of Mathematicians in Moscow. Reports to the Commission by the Conference Board of the Mathematical Sciences (1966) in the United States and by the Association of Teachers of Mathematics (1966) in England highlighted the importance of problems in mathematics instruction and indicated that educators need to know much more about using problems to stimulate independent and creative thinking.

The Cambridge Conference on School Mathematics (Educational Services Incorporated, 1963) urged curriculum developers to devote more time and energy to the creation of problem sequences, with special emphasis on problems that can be used to introduce new mathematical ideas. If this suggestion is taken seriously, and when complex, challenging mathematical problems assume a more central role in the curriculum, more research on such problems will be needed.

Problem solving is not now being investigated systematically by mathematics educators. Few studies build on previous research; few studies have an explicit theoretical rationale. There are signs, however, that some mathematics educators are beginning to borrow ideas and techniques from recent psychological studies of higher cognitive processes. As more educational researchers appear who are trained both in mathematics and in psychology, research on problem solving in mathematics may attain a direction and cohesiveness it currently lacks.

---

\*Dr. James W. Wilson, University of Georgia, served as consultant to Dr. Kilpatrick on the preparation of this chapter. Richard Pocock assisted in the initial search of the literature.



## *Bibliographies and Reviews*

Problem solving in elementary school mathematics was the subject of several recent reviews. Riedesel (1969) listed 83 outstanding articles and research reports published during the past fifty years and noted some obvious implications of this research for arithmetic teachers. Gorman (1967), in a more systematic and critical analysis, identified 293 studies on word problems conducted between 1925 and 1965. Only 37 of these studies, mostly doctoral theses, were deemed "acceptable" according to Gorman's criterion of high internal validity. A monumental survey by Suydam (1967) of published research on elementary school mathematics from 1900 to 1965 yielded 84 studies of problem solving. Both Suydam and Gorman found conflicting results and a generally low quality of design and reporting.

### *Problem-Solving Ability*

Without necessarily conceiving of the ability to solve mathematical problems as a unitary phenomenon, one can learn something of its nature by examining the relationships between an individual's success in problem solving and other characteristics of his thinking and personality. Studies of problem-solving ability range from straightforward comparisons of group performance to intricate factor analyses.

Tate and Stanier (1964) analyzed the performance of good and poor problem solvers on tests of critical thinking and practical judgment. Subjects were 234 junior high school students whose scores on a composite measure of problem-solving ability, including mathematical and quantitative reasoning problems, deviated markedly from a regression line of problem solving on IQ. On the critical thinking tests, the poor problem solvers tended to avoid the judgment "not enough facts" and to make unqualified "true" or "false" judgments. On the practical judgment test, they tended to select answers having a high affective component. Tate and Stanier argued that the errors might stem from response sets having a temperamental rather than an intellectual basis.

Success in solving word problems in mathematics obviously depends upon skills in reading and computation, but the relative contribution of these skills is not clear. Martin (1963) found that each of the factors of reading comprehension, computation, abstract verbal reasoning, and arithmetic concepts was correlated with problem solving as measured by the Arithmetic Problem-solving Test of the Iowa Tests of Basic Skills given to fourth and eighth graders. The practical correlation between reading and problem solving with computation held constant (about .5) was higher at both grade levels than the partial correlation between computation and problem solving with reading held constant (about .4). As Martin

suggested, the relationship between problem-solving ability and its underlying skills, particularly higher-order verbal skills, is probably more complex than has been supposed.

Two factor analytic studies of problem solving in mathematics were synthesized by Werdelin (1966), who rotated the two factor matrices to a congruent structure. The loadings on the five factors isolated in each study were virtually identical. Tests of problem solving loaded most strongly on a General Reasoning factor and to a lesser extent on a Deductive Reasoning and a Numerical factor. The other factors, Space and Verbal Comprehension, were unrelated to problem solving.

*Additional References:* Balow (1964); Dye and Very (1968); Koopman (1964); Sheehan (1968); Very (1967).

## *Problem-Solving Tasks*

### *Problem Content*

A perennial issue in mathematics education concerns the use of problems that are closely related to students' interests and experiences. Travers (1967) asked 240 male high school freshmen to choose and solve one of two problems that were identical in structure (numbers used, operations required, etc.) but different in setting. The subjects showed strong preferences for "social-economic" situations (e.g., selling hot dogs) compared with "mechanical-scientific" situations (e.g., testing spark plugs) and "abstract" situations (e.g., solving secret codes). The last situations were particularly unpopular. General mathematics students showed stronger differential preferences than algebra students, and there were some tendencies, although slight, for problem preferences to be related to vocational interests as expressed on the Kuder Preference Record.

Scott and Lighthall (1967) tested the hypothesis that disadvantaged children would perform relatively better on problems whose content dealt with lower needs, such as food and shelter, than on problems whose content dealt with higher needs, such as mastery and education. Need content of the problems was not related to degree of disadvantage of third and fourth graders. A principal components analysis of the data suggested that factors associated with the difficulty and the mathematical content of the items, rather than the need content, accounted for differences in performance.

### *Problem Structure*

Steffe (1967) investigated the effects of two variations of the language used in a problem on its difficulty. Twenty one-step addition problems were presented orally to ninety first graders in individual interviews. In

ten of the problems the names for the two sets to be combined and the total were the same ("There are four cookies on one plate and two cookies on another plate. How many cookies are on the plates?"), and in ten of the problems the names for the three sets were different ("Mary has four kittens and two goldfish. How many pets does Mary have?"). Half of the subjects were given problems in which an existential quantifier was used at the beginning of the problem ("There are some cookies on two plates"), and half were given problems without the quantifier. The presence of the quantifier had no effect on problem difficulty, but the problems with a common name for the sets proved to be significantly easier than the problems with different names for the sets.

Thompson (1967) reported that the effects of readability and mental ability on arithmetic problem-solving performances were interactive. Although ease of reading was associated with higher performance at both high and low levels of mental ability, the effect was greater with subjects of low mental ability.

The ordering of the data in a problem is another factor in problem difficulty. Although it is apparently immaterial whether the question is placed at the beginning or end of the problem statement (Williams and McCreight, 1965), a problem can be made more difficult, especially for low ability subjects, by presenting the data in some order other than that used to solve the problem (Burns and Yonally, 1964).

The sequence in which problems are presented is also important. Suppes, Loftus, and Jerman (1969) found that structural variables such as the number of different operations needed to solve a problem were less influential in determining its difficulty than whether or not the problem could be solved the same way as the preceding one.

### *Problem-Solving Processes*

Since the solution of a problem—a mathematics problem in particular—is typically a poor index of the processes used to arrive at that solution, problem-solving processes must be studied by getting subjects to generate observable sequences of behavior. Psychologists have devised numerous techniques for studying problem solving in the laboratory, but mathematical problems are seldom used in such research. One of the few psychologists to study how complex mathematical problems are solved was Karl Duncker (1945), who used the "thinking aloud" technique. Thinking aloud, although out of favor for years, has recently reappeared with the advent of information-processing approaches to the study of problem solving (see Hunt, 1968), and mathematics educators have begun to make use of the technique in their studies.

Information-processing theories have also stimulated interest in the

role of heuristic rules in problem solving. The eminent mathematician George Polya (1957, 1962, 1965), in his books for teachers, set forth maxims for problem solving which, he postulated, correspond to mental actions. The most impressive evidence for the validity of Polya's observations on the problem-solving process has come from work on computer simulation of human behavior. Programmers have found that the incorporation of general heuristic rules, such as working backward or using a diagram, not only facilitates problem solving, but also results in performance by the computer that closely resembles the behavior of humans struggling with similar problems.

Paige and Simon (1966) compared the protocols of subjects asked to think aloud as they solved algebra word problems with the processes used in a computer program for translating English sentences into equations and then solving them. Analysis of the protocols showed that subjects used some kind of internal representation of the physical situation described by the problem in framing their equations. When given "contradictory" problems in which equations can be written even though the solution is physically impossible, subjects consistently differed in their ability to detect the contradiction. Paige and Simon concluded that good problem solvers are more likely than poor problem solvers to discover contradictions of this sort. Their observation, though confirmed by Kennedy, Eliot, and Krulee (in press), conflicts with Krutetskii's (1969) finding that for some contradictory problems capable secondary school students and adults make more mistakes than less capable subjects, who use a concrete interpretation and thereby discover the fallacy. The difference in results may be attributable to the Soviet educational practice of teaching mathematics problems as representative of certain "types." Krutetskii's more capable subjects apparently saw the contradictory problems as embodying a type, recalled the model solution, and then mechanically substituted the (illogical) data into the model.

*Additional References:* Cunningham (1966); Kilpatrick (1967).

### *Instructional Programs*

Recent years have witnessed the development of increasingly sophisticated theory-based programs of instruction in problem solving. Though most studies continue to be evaluations of a single device or technique, some attempts have been made to develop broader programs having an explicit theoretical rationale.

#### *Training in Heuristic Methods*

Covington and Crutchfield (1965) reported several studies with the *General Problem Solving Program (GPSP)*, a well-conceived and appar-



ently successful program they devised for teaching children to apply heuristic strategies to problems. Though the problems are not mathematical, the strategies are appropriate to mathematical problem solving. The program consists of a series of self-instructional booklets that give the pupil repeated experiences in solving interesting problems and show him such strategies as planning one's attack, searching for uncommon ideas, transforming the problem, and using analogies. Covington and Crutchfield found dramatic gains for an instructional group of fifth and sixth graders as compared with a control group on attitude inventories and tests of problem-solving ability and creative thinking. Five months after instruction, gains in problem-solving ability diminished somewhat but were still statistically significant; gains in creative thinking had become marginal (Covington, 1968).

Concerned about several apparent failures to replicate Covington and Crutchfield's findings (Alton et al., 1967; Ripple and Dacey, 1967) and curious about the question of nonspecific transfer, Treffinger and Ripple (1968) investigated the effectiveness of the GPSP on verbal creativity, general problem solving, arithmetic problem solving, and attitudes at each of the grades 4 through 7. Only a few differences on the verbal creativity and problem-solving tests—not much more than would be expected by chance—reached statistical significance, and these differences formed no obvious pattern. (The results for arithmetic problem solving may have been clouded by the difficulty and the low reliability of the tests used.) Statistically significant differences favoring the instructed group were found on a measure of general attitudes about creative thinking and problem solving at all four grade levels. The results suggested that although the GPSP may be successful in promoting some transfer to novel problems, unless the format of the problems resembles that of the training materials, transfer is likely to be minimal.

Attempts to promote transfer of heuristics from miscellaneous training tasks to problems from disciplines such as mathematics raise questions as to how general or how specific the heuristics should be. James Wilson (1967) predicted that subjects taught specific heuristics would perform better on training tasks but worse on transfer tasks than would subjects taught general heuristics. Subjects studied self-instructional booklets on two theorem-proving tasks, one in symbolic logic and the other in elementary algebra. For each task, subjects were taught to use one of three kinds of heuristic: task-specific (applicable to the training task only), means-end (locating the key difference between the given situation and the goal and then searching for a means to reduce the difference), and planning (omitting details in the given situation and working out a proposed solution in general terms). A  $3 \times 3 \times 2$  factorial design was used, with three levels of heuristic for each task and two orders of task presenta-

tion. Dependent variables were derived from performance on the training tasks and five transfer tasks (two similar and three dissimilar in format to the training tasks). Task-specific heuristics did not facilitate performance on the training tasks; in fact, on one training task the planning heuristic was superior to the others. On the dissimilar transfer task, the planning heuristic was superior to the others; otherwise there were no significant main effects. Significant interactions suggested that a combination of heuristics during training facilitated performance on some of the transfer tasks and that general heuristics learned in the first training task were practiced on the second task, thereby facilitating transfer.

Studies of training programs in heuristic methods are hampered because we know very little about how subjects use heuristics and nothing at all about how they adapt heuristics to different kinds of problems. However, the success of these training programs in producing nonspecific transfer under certain conditions is certain to encourage further research on heuristics.

*Additional References:* Brian (1966); Post (1967).

### *Learning by Discovery*

Wills (1967) investigated the effect of learning by discovery on problem-solving ability. Two groups of eight intermediate algebra classes were given two weeks' instruction on figurate numbers and recursive definitions. The instructional materials, in a workbook format, introduced a topic by presenting a difficult problem that required a generalization, guiding the student with a series of simpler problems, prompting the student to look for a pattern in the problems, and giving the student a check on his generalization. In one group of classes, the teachers discussed various heuristic methods for discovering the generalization in the instructional materials; in the other group, the teachers gave no such guidance. Before and after instruction, both groups took a sixty-item test on mathematical topics not covered in the unit. On the posttest both groups doubled their pretest performance, whereas a control group that took the tests without the intervening instruction made only a minor gain. Guidance by the teacher on heuristic methods apparently did not contribute to the gains; the adjusted means of the two experimental groups on the posttest did not differ significantly.

*Additional Reference:* Scandura (1964).

### *Other Methods*

Much attention has been focused in recent years on finding methods and devices that would improve problem solving without putting the child

in the kind of straightjacket provided by formal analysis and other prescriptive techniques. The traditional approach to the solving of word problems in the elementary grades has been characterized as a "wanted-given" procedure—the child is taught to ask himself, "What is wanted?" and "What is given?" and then to perform the appropriate operations on the data to yield values for the unknowns.

John Wilson (1967) contrasted one version of the wanted-given approach, in which the child analyzes structural relationships between the data and the unknowns, with an "action-sequence" approach, in which the child looks for the operations suggested by the sequence of actions in the problem. In both approaches the child writes and solves a number sentence that expresses the structure of the problem. Fifty-four fourth grade subjects at three levels of mental ability were randomly assigned to a wanted-given, an action-sequence, and a control group. Each group was given three periods of instruction a week for nine weeks, using worksheets supplemented by instruction from the teacher (except for the control group, which just worked the problems and then spent the remaining time on other activities). A  $3 \times 3$  factorial analysis of variance followed by multiple comparison tests showed that on measures of ability to choose correct operations, ability to solve problems, and speed in solving problems, the wanted-given group performed better than the other two groups after three, six, and nine weeks of instruction, and again nine weeks after instruction had ended. Wilson's study provides strong support for the superiority of the wanted-given approach, although as Zweng (1968) observed, Wilson's wanted-given treatment varies considerably from the traditional one.

*Additional References:* Bechtold (1965); Denmark (1964); Gangler (1967); Keil (1964); Koch (1965); Lerch and Hamilton (1966); Riedesel (1964); Stuart (1965); VanderLinde (1964).

### *Teacher Influences*

Polya (1962, 1965) argued that teachers cannot teach problem solving unless they have had some problem-solving experience themselves. Though this argument has not been tested directly, Godgart (1964) demonstrated that, at least on one measure, teachers' problem-solving ability in mathematics is not related to pupil progress. The mathematics test of the Sequential Tests of Educational Progress (STEP) was administered to 35 fourth-grade teachers, and the Arithmetic Problem-Solving subtest of the Iowa Tests of Basic Skills (ITBS) was administered twice to their pupils—at the beginning of the fourth grade and again at the beginning of the fifth grade. When the teachers were divided into five equal groups according to their performance on the STEP, and analysis of covariance was performed on class means with the ITBS pretest as the covariate and the

ITBS posttest as the dependent variable, the groups did not differ significantly. Furthermore, teachers' problem-solving ability as measured by the STEP was unrelated to such background measures as age, tenure status, undergraduate major, and number of mathematics content and methods courses taken.

If the teacher does have an impact on pupils' problem-solving ability, the focus of this impact must be the classroom. Stilwell (1967) adapted the Flanders interaction-analysis scheme to study problem-solving activity in geometry classrooms. One of Stilwell's most interesting findings concerned the relatively small amount of class time (less than 3 per cent of all problem-solving activity) spent in discussing a method for solving a problem. Looking back at a problem or ahead to its implications occupied 7 per cent of the problem-solving activity, with teacher requests to look back at the solution correlated positively and significantly with greater teaching experience.

### Conclusion

A good share of the research in mathematics education, now as in the past, is being done by doctoral students. Though many theses on problem solving in mathematics are of a quality and sophistication that surpasses the general level of journal articles on this topic, the theses are relatively inaccessible. The forthcoming *Journal for Research in Mathematics Education* should provide a forum for dissertation research that has heretofore gone unpublished and unpublicized.

As research in mathematics education becomes more sophisticated in design, more explicitly grounded in theory, and more closely allied to developments in other fields, the mathematics educators' one-shot comparisons of ill-defined "methods" and the psychologists' laboratory studies of arbitrary, highly artificial problems should give way to diagnostic, long-term studies of learning and thinking in school settings. Evidence of this trend can be noted in the studies reviewed, as can a general movement toward more complex designs and analyses.

Unfortunately, the increasing complexity of design has been accompanied by an increasing number of methodological blunders, such as the inappropriate use of analysis of covariance and the use of subjects as experimental units when intact classes have been assigned to treatments. More disturbing still is the investigators' apparent ignorance that statistical assumptions are being violated. Mathematics educators, of all people, should be highly skeptical about the congruence between an analytical model and the "real-world" data.

Much has been said lately about the need for large-scale, complex studies in mathematics education, but the researcher—most likely a doc-



toral student—who chooses to investigate problem solving in mathematics is probably best advised to undertake clinical studies of individual subjects (children gifted in mathematics, children for whom mathematics is particularly difficult, etc.), not only because clinical studies are more commensurate with limited financial and administrative resources, but also because our ignorance in this area demands clinical studies as precursors to larger efforts.

### Bibliography

- Association of Teachers of Mathematics, Research and Development Panel. *The Development of Mathematical Activity in Children: The Place of the Problem in This Development*. Nelson, Lancashire, England: The Association, Vine Street Chambers, Nov. 1966.
- Burns, Paul C. and Yonally, James L. Does the Order of Presentation of Numerical Data in Multi-Steps Arithmetic Problems Affect Their Difficulty? *School Science and Mathematics* 64: 267-70; 1964.
- Conference Board of the Mathematical Sciences. *The Role of Axiomatics and Problem Solving in Mathematics*. Boston: Ginn, 1966.
- Covington, Martin V. Promoting Creative Thinking in the Classroom. *Research and Development Toward the Improvement of Education*. (Edited by Herbert J. Klausmeier and George T. O'Hearn.) Madison, Wis.: Dembar Educational Research Services, 1968. Chapter 3, pp. 22-30.
- Covington, Martin V. and Crutchfield, Richard S. Facilitation of Creative Problem Solving. *Programed Instruction* 4, No. 4: 3-5, 10; 1965.
- Duncker, Karl. On Problem-Solving. *Psychological Monographs: General and Applied*, Vol. 58, No. 5 (Whole No. 270). Washington, D. C.: American Psychological Assoc., 1945.
- Educational Services Incorporated. *Goals for School Mathematics*. Report of the Cambridge Conference on School Mathematics. Boston: Houghton Mifflin, 1963.
- Godart, Martin David. *The Arithmetic Problem-Solving Ability of Teachers as Related to Their Effectiveness in Teaching Problem-Solving*. Doctor's thesis. Storrs: Univ. of Conn., 1964. Abstract: *Dissertation Abstracts* 26: 793-94; No. 2, 1965.
- Gorman, Charles J. *A Critical Analysis of Research on Written Problems in Elementary School Mathematics*. Doctor's thesis. Pittsburgh: Univ. of Pittsburgh, 1967. Abstract: *Dissertation Abstracts* 28: 4818A-19A; No. 12, 1968.
- Hunt, Earl. Computer Simulation: Artificial Intelligence Studies and Their Relevance to Psychology. *Annual Review of Psychology* 19: 135-68; 1968.
- Kennedy, George; Eliot, John; and Krulee, Gilbert. *The Analysis of Error Patterns. Psychology in the Schools*, in press.
- Krutetskii, V. A. An Investigation of Mathematical Abilities in Schoolchildren. *The Structure of Mathematical Abilities*. Soviet Studies in the Psychology of Learning and Teaching Mathematics, Vol. 2. (Edited by Jeremy Kilpatrick and Izaak Wirszup.) Stanford: School Mathematics Study Group, 1969. Pp. 5-57.
- Martin, Mavis Doughty. *Reading Comprehension, Abstract Verbal Reasoning, and Computation as Factors in Arithmetic Problem Solving*. Doctor's thesis. Iowa City: State Univ. of Iowa, 1963. Abstract: *Dissertation Abstracts* 24: 4547-48; No. 11, 1964.
- Olton, Robert M. et al. *The Development of Productive Thinking Skills in Fifth-Grade Children*. Technical Report No. 34. Madison: Wis. Research and Development Center for Cognitive Learning, Univ. of Wis., Nov. 1967.
- Paige, Jeffery M. and Simon, Herbert A. Cognitive Processes in Solving Algebra Word Problems. *Problem Solving: Research, Method, and Theory*. (Edited by Benjamin Kleinmuntz.) New York: John Wiley, 1966. Chapter 3, pp. 51-119.
- Polya, G. *How to Solve It*. Second edition. Garden City, N. Y.: Doubleday, 1957.
- Polya, George. *Mathematical Discovery: On Understanding, Learning, and Teaching Problem Solving*. New York: John Wiley, 1962, Vol. 1; 1965, Vol. 2.

- Riedesel, C. Alan. Problem Solving: Some Suggestions from Research. *Arithmetic Teacher* 16: 54-58; 1969.
- Ripple, Richard E. and Dacey, John. The Facilitation of Problem Solving and Verbal Creativity by Exposure to Programed Instruction. *Psychology in the Schools* 4: 240-45; 1967.
- Scott, Ralph and Lighthall, Frederick F. Relationship Between Content, Sex, Grade, and Degree of Disadvantage in Arithmetic Problem Solving. *Journal of School Psychology* 6: 61-67; 1967.
- Steffe, Leslie P. *The Effects of Two Variables on the Problem-Solving Abilities of First-Grade Children*. Technical Report No. 21. Madison: Wis. Research and Development Center for Cognitive Learning, Univ of Wis., Mar. 1967.
- Stilwell, Merle Eugene. *The Development and Analysis of a Category System for Systematic Observation of Teacher-Pupil Interaction During Geometry Problem-Solving Activity*. Doctor's thesis. Ithaca, N. Y.: Cornell Univ., 1967. Abstract: *Dissertation Abstracts* 28: 3083A; No. 8, 1968.
- Suppes, Patrick; Loftus, Elizabeth F.; and Jerman, Max. *Problem-Solving on a Computer-Based Teletype*. Technical Report No. 141, Psychology Series. Stanford: Institute for Mathematical Studies in the Social Sciences, Stanford Univ., Mar. 1969. (Mimeo.)
- Suydam, Marilyn N. The Status of Research on Elementary School Mathematics. *Arithmetic Teacher* 14: 684-89; 1967.
- Tate, Merle W. and Stanier, Barbara. Errors in Judgment of Good and Poor Problems Solvers. *Journal of Experimental Education* 32: 371-76; Summer 1964.
- Thompson, Elton Noel. *Readability and Accessory Remarks: Factors in Problem Solving in Arithmetic*. Doctor's thesis. Stanford: Stanford Univ., 1967. Abstract: *Dissertation Abstracts* 28: 2464A-65A; No. 7, 1968.
- Travers, Kenneth J. A Test of Pupil Preference for Problem-Solving Situations in Junior High School Mathematics. *Journal of Experimental Education* 35: 9-18; Summer 1967.
- Treffinger, Donald J. and Ripple, Richard E. *The Effects of Programed Instruction in Productive Thinking on Verbal Creativity and Problem Solving Among Elementary School Pupils*. Cooperative Research Project No. 8-B-002. Ithaca, N. Y.: Cornell Univ., 1968.
- Werdelin, I. *A Synthesis of Two Factor Analyses of Problem Solving in Mathematics*. Didakometary, No. 8. Malmö, Sweden: Malmö School of Education, 1966.
- Williams, Mary Heard and McCreight, Russell W. Shall We Move the Question? *Arithmetic Teacher* 12: 418-21; 1965.
- Wills, Herbert, III. *Transfer of Problem Solving Ability Gained Through Learning by Discovery*. Doctor's thesis. Urbana: Univ. of Ill., 1967. Abstract: *Dissertation Abstracts* 28: 1319A-20A; No. 4, 1967.
- Wilson, James William. *Generality of Heuristics as an Instructional Variable*. Doctor's thesis. Stanford: Stanford Univ., 1967. Abstract: *Dissertation Abstracts* 28: 2575A; No. 7, 1968.
- Wilson, John W. The Role of Structure in Verbal Problem Solving. *Arithmetic Teacher* 14: 486-97; 1967.
- Zweng, Marilyn J. A Reaction to "The Role of Structure in Verbal Problem Solving." *Arithmetic Teacher* 15: 251-53; 1968.

### Additional References

- Balow, Irving H. Reading and Computation Ability as Determinants of Problem Solving. *Arithmetic Teacher* 11: 18-22; 1964.
- Bechtold, Charles August. *The Use of Extraneous Material in Developing Problem-Solving Ability of Pupils in Algebra I*. Doctor's thesis. New York: Columbia Univ., 1965. Abstract: *Dissertation Abstracts* 26: 3105; No. 6, 1965.
- Brian, Richard Boring. *Processes of Mathematics: A Definitional Development and an Experimental Investigation of Their Relationship to Mathematical Problem Solving Behavior*. Doctor's thesis. College Park: Univ. of Md., 1966. Abstract: *Dissertation Abstracts* 28: 1202A; No. 4, 1967.

- Cunningham, John D. Rigidity in Children's Problem Solving. *School Science and Mathematics* 66: 377-89; 1966.
- Denmark, Ewell Thomas, Jr. A Comparative Study of Two Methods of Teaching Elementary Algebra Students to Use the Algebraic Technique to Solve Verbal Problems. Doctor's thesis. Tallahassee: Fla. State Univ., 1964. Abstract: *Dissertation Abstracts* 25: 5295-96; No. 9, 1965.
- Dye, Neal W. and Very, Philip S. Growth Changes in Factorial Structure by Age and Sex. *Genetic Psychology Monographs* 78: 55-88; 1968.
- Gangler, Joseph Michael, Jr. An Experimental Study of the Effects of Participation and Motivation on the Problem Solving Ability of College Freshmen. Doctor's thesis. New York: Columbia Univ., 1967. Abstract: *Dissertation Abstracts* 28: 21578; No. 5, 1967.
- Keil, Gloria Emilie. Writing and Solving Original Problems as a Means of Improving Verbal Arithmetic Problem Solving Ability. Doctor's thesis. Bloomington: Ind. Univ., 1964. Abstract: *Dissertation Abstracts* 25: 7109-10; No. 12, 1965.
- Kilpatrick, Jeremy. Analyzing the Solution of Word Problems in Mathematics: An Exploratory Study. Doctor's thesis. Stanford: Stanford Univ., 1967. Abstract: *Dissertation Abstracts* 28: 4380A; No. 11, 1968.
- Koch, Elmer A., Jr. Homework in Arithmetic. *Arithmetic Teacher* 12: 9-13; 1965.
- Koopman, Norbert Earl. Evaluations by Superior High School Students of Their Problem Solving Performances. Doctor's thesis. Madison: Univ. of Wis., 1964. Abstract: *Dissertation Abstracts* 25: 3398; No. 6, 1964.
- Lerch, Harold H. and Hamilton, Helen. A Comparison of a Structured-Equation Approach to Problem Solving With a Traditional Approach. *School Science and Mathematics* 66: 241-46; 1966.
- Post, Thomas Robert. The Effects of the Presentation of a Structure of the Problem-Solving Process Upon Problem-Solving Ability in Seventh Grade Mathematics. Doctor's thesis. Bloomington: Ind. Univ., 1967. Abstract: *Dissertation Abstracts* 28: 4545A; No. 11, 1968.
- Riedesel, C. Alan. Verbal Problem Solving: Suggestions for Improving Instruction. *Arithmetic Teacher* 11: 312-16; 1964.
- Scandura, Joseph M. An Analysis of Exposition and Discovery Modes of Problem Solving Instruction. *Journal of Experimental Education* 33: 149-59; Winter 1964.
- Sheehan, T. Joseph. Patterns of Sex Differences in Learning Mathematical Problem-Solving. *Journal of Experimental Education* 36: 84-87; Summer 1968.
- Stuart, Alvin James. Effects Upon Pupil Performances in Arithmetic of Instructional Programs Differing in Amounts of Emphasis upon Computational Structure and Verbal Problem Solving. Doctor's thesis. Athens: Ohio Univ., 1965. Abstract: *Dissertation Abstracts* 27: 4058A; No. 12, 1967.
- VanderLinde, Louis F. Does the Study of Quantitative Vocabulary Improve Problem-Solving? *Elementary School Journal* 65: 143-52; 1964.
- Very, Philip S. Differential Factor Structures in Mathematical Ability. *Genetic Psychology Monographs* 75: 169-207; 1967.

## AUTHOR

JEREMY KILPATRICK Address: Teachers College, Columbia University, New York, New York Title: Assistant Professor of Mathematics Age: 34 Degrees: A.B. and M.A., Univ. of California; M.S. and Ph.D., Stanford Univ. Specialization: Problem Solving in Mathematics.

## 11: CLASSROOM TEACHING OF MATHEMATICS

James Fey\*

*University of Maryland*

What is a good mathematics teacher? What are his personal characteristics? What sort of preservice and inservice education is most appropriate for him? What teaching method is most effective in mathematics classes? Few questions in mathematics education are more important than these, and few generate more vigorous debate or consistent disagreement. Despite long and active interest in the problems, research offers few important guidelines in the search for personal attributes, classroom styles, or educational preparation of successful teachers. Extensive investigation has failed to show significant or consistent correlation between fundamental characteristics of teachers—such as experience, knowledge of mathematics, collegiate preparation, or attitudes toward mathematics—and the achievement or attitudes of their students. Evidence from comparisons of two or more teaching methods supports no one method as superior in mathematics teaching.

Much current research in mathematics teaching continues the traditional search for the elusive "good teacher." However, there is growing evidence of creative, yet careful, new research strategies and techniques.

This chapter covers research in mathematics teaching reported from 1964 through 1968. Significant results and promising direction for study are indicated; this chapter is not a catalog of research. Following a discussion of the quality of current research, the review is divided into two major areas: 1) personal characteristics of teachers and 2) teacher classroom behavior, including methods of teaching. The quality and originality of recent research into characteristics and behavior of effective teachers were not evident in studies of preservice and inservice teacher education. Because of this weakness and space limitations, research in teacher education has not been included in this issue of the *Review*.

### *Weaknesses and Strengths of Current Research*

During the period reviewed, many of the attempts to identify critical

---

\*Drs. F. Joe Crosswhite, Ohio State University, and H. Laverne Thomas, State University of New York College at Oneonta, served as consultants to Dr. Fey on the preparation of this chapter.



personal characteristics of teachers were limited by narrow vision of what the key variables might be and by oversimplified measures of effectiveness. But in several studies, investigators uncovered previously unexamined teacher attributes that appear to have promise in predicting effectiveness. Other investigators attempted to broaden the accepted criteria of effectiveness; instead of a simple *good* or *bad* appraisal of effectiveness based on standardized student achievement tests, they sought a detailed description of student outcomes—cognitive and affective, manifest and latent.

Inadequacy of criterion measures also plagued research in methods of teaching mathematics. Instead of describing the outcomes associated with particular instructional styles, methods research was usually an attempt to show the superiority of one *name* method over another, as measured by achievement tests. This approach is insensitive to the differential effects that each individual treatment might have. As a consequence, results from methods were inconclusive; the strong and weak features of each method tended to counterbalance each other. The most promising research in mathematics teaching methods focused on the crucial problem of aptitude-treatment interaction (ATI) in an attempt to discern the unique impact of various teaching styles.

Methods research also suffered because most studies were small-scale projects; often a single investigator taught one or two experimental classes and a like number of control classes and assumed responsibility for evaluation of the results. Thus, even when significant results were reported, the impact of such findings on teaching practice was minimal. The investigator's personal involvement in the research made bias inevitable, and the methods compared were seldom described in behavioral terms precise enough to allow repetition of the methods in other classes with other teachers. However, some research, with the benefit of sophisticated technology for recording classroom activity, made a strong bid to sharpen definitions of teaching methods and, thus, make methods research more reliable and generalizable.

New approaches to the study of mathematics teaching were fresh steps into what has been an extremely discouraging area of research. An equally important development was the growth of interest in a theory of mathematics instruction. Even if the various investigation of mathematics teaching are successful in describing relationships among pairs of key instructional variables, one important task remains—carrying these findings into widespread classroom practice. Implementation of teaching research requires integration of all the relevant knowledge into a coherent prescription for instructional behavior. Tentative theories of mathematics instruction have begun to provide a conceptual framework indicating important research questions and the relationships of widely varied research findings to classroom teaching.

In the past, this translation of scattered research results into teaching practice was left largely to the ingenuity of individual teachers in the classroom. Thus, Henderson (1963, p. 1025) concluded that as of 1963, mathematics teaching was an art, not a science and "for it to move toward becoming a science, we need much more empirical research to test current theories. But we also need new theoretical concepts or orientations that will provoke different questions to be asked." Research in mathematics teaching since that date has shown definite signs of meeting his demands for rigorous empiricism and new theoretical organization.

Much of the credit for new theoretical and empirical vitality in research is due to the appearance of Gage's *Handbook of Research in Teaching* (1963). The influence of research guidelines in the *Handbook* was not noticeable in early research covered by this review, but recent studies were increasingly well designed and administered.

### *Personal Characteristics of Teachers*

Although previous investigations of teacher characteristics and effectiveness failed to yield a consistent or definitive profile of the effective mathematics teacher, introduction of strikingly new mathematics curricula provoked renewed interest in teacher characteristic research during the past ten years. The new curricula brought new content, greater mathematical precision, and a concept (rather than a skill) orientation to mathematics study and teaching. These changes reopened several fundamental questions: 1) Is there an identifiable relationship between teacher knowledge of mathematics and student achievement? 2) Do teacher age or years of experience influence student achievement? 3) How do teacher attitudes toward the new mathematics affect student attitudes or achievement?

#### *Minnesota National Laboratory Study*

The most comprehensive, sophisticated, and imaginative investigation of teacher characteristics and student achievement was conducted by the Minnesota National Laboratory for the Improvement of Secondary Mathematics as part of its effort to evaluate experimental curricular materials of the School Mathematics Study Group (SMSG). SMSG wanted to determine whether the new curricular materials it produced were suitable for widespread implementation by teachers of varying ability and qualifications.

An exploratory study involving 21 teachers of grades 6-8 during the 1958-1959 school year and one involving 127 teachers of grades 7-12 during the 1959-1960 school year failed to reveal any significant correlation between students' achievement and the experience, collegiate courses

or grades, and professional activity of their teachers. However, analysis of daily logs kept by the participating teachers during the 1958-1959 study and of monthly checklist reports made during 1959-1960 indicated that the most effective and least effective teachers could be differentiated by measures of productive thinking—the ability to generate ideas about success or failure of various teaching lessons and ideas about alternative procedures for teaching particular concepts.

These preliminary investigations led to design of the main study (Torrance and Parent, 1966) in which further evidence was sought on the correlates of teaching effectiveness that might be described as qualifications to teach—experience, collegiate preparation, and professional activity—and the importance of variables in teacher-pupil interaction and teacher productive thinking ability. During the 1960-1961 and the 1961-1962 school years, each of 63 teachers taught the experimental SMSG course to one class.

Results from the 1960-1961 and the 1961-1962 studies confirmed the suspicion that effectiveness of teachers using the SMSG materials was not significantly correlated with teachers' experience, collegiate courses and grades, or participation in professional activities. Most and least effective teachers were not differentiated by the amount of time they spent in preparation for teaching. There was only a weak indication that procedures in making assignments, explaining new material, conducting learning and thinking experiences relevant to previously assigned material, and evaluating and responding to student performance made a difference in teacher effectiveness. Application of Flanders's system of Interaction Analysis failed to locate significant differences in the patterns of classroom behavior developed by effective and ineffective teachers.

The best predictor of teaching effectiveness was the productive thinking ability of teachers. The most effective teachers produced more ideas about indications of success or failure in their teaching, causes of success or failure, and alternate ways to teach course concepts. The student attitude and learning activities questionnaires revealed a number of significant indicators of teaching effectiveness. This suggests that teacher effectiveness is intimately connected with pupil attitudes and perceptions about the methods of their teachers, the school, text materials, and the class as a group.

Thus, the Minnesota study confirmed earlier indications that the search for predictors of effective teaching must move beyond the gross measures of ability and formally identifiable qualifications. Apparently, qualification beyond some minimal standard is not reflected in greater effectiveness. In the search for more subtle but crucial factors in teaching success, the Minnesota Study provided fresh insight into topics and techniques for future research.

### *Secondary Mathematics Teachers*

One of the teacher variables measured indirectly in the Minnesota Study was the teachers' knowledge of contemporary mathematics—a factor that would logically seem correlated with effectiveness. Massie (1967) asked a panel of 12 mathematics educators to identify the characteristics of contemporary mathematics; he then developed and pilot-tested an examination called *Contemporary Mathematics: A Test for Teachers*. Drawing a sample of 273 prospective teachers in colleges of eight states and 58 teachers with experience in modern mathematics, he also gathered some descriptive data about the mathematical competence of the chosen sample. Lyng (1968) used Massie's test to uncover correlations between teacher knowledge and a variety of other attributes, but the results offer no insight into the question of teacher effectiveness.

Kennedy (1963) developed a test of skill in solving mathematics teaching problems. The test consisted of 17 tape-recorded problem situations excerpted from actual algebra classes, each of which was followed by oral questions about solutions. A panel of experts devised rating scales for responses. The test was administered to 311 teachers of varying backgrounds. Scores on the test differentiated groups of pre- and inservice teachers at the .01 level of significance in the following order from poorest to best: 1) non-mathematics majors, 2) elementary education majors, 3) preservice mathematics teachers without methods course experience, 4) preservice mathematics teachers who had a methods course, 5) preservice mathematics teachers who had practice teaching experience and a methods course, and 6) experienced mathematics teachers. But again, there was no way to consider the relationship between differences on the test and teaching effectiveness.

### *Elementary Mathematics Teachers*

The initial thrust of recent curriculum innovation was in the secondary school, but emphasis soon shifted to the elementary school. Elementary school teachers, already faced with the task of specializing in every subject of the curriculum, were asked to learn a great deal of new mathematics and to learn and teach the traditional content from a new point of view. Concern for teachers' subject-matter competence made teacher ability a popular area for research.

Several studies attempted to assess the mathematical competence of preservice and inservice elementary school teachers. The results consistently showed that teachers do not have the knowledge of modern mathematics considered to be a prerequisite to effective teaching. For example, in a penetrating study of prospective teachers in Negro private colleges, Carroll (1964) compared normative data (STEP tests) with the mathematical



knowledge of 358 teachers just before graduation from college; he analyzed the particular strengths and weaknesses in the knowledge of these teachers. He found that the group's functional competence in mathematics was at the level of seventh and eighth grade students, with knowledge of geometry and probability the weakest.

The grim picture of mathematical competence among elementary teachers, as reported during the 1964-1968 period, continued an historic pattern. But it also indicated slow awakening in colleges to the needs of teachers faced with a curriculum vastly different from the one they experienced in their own schooling. In the elementary school, as in the high school, the relationship between teacher mathematical competence and student achievement remained unclear. The available information was not encouraging.

For example, Moore (1965) examined the hypothesis that there is a positive relation between level of teacher understanding (as measured on the Glennon Test of Mathematical Understanding) and student achievement (as measured by the SRA Arithmetic Series). Using 10 fourth grade and 11 sixth grade classes as subjects, Moore found no significant correlation between mean classroom gain in achievement and teacher understanding; classes of more knowledgeable teachers showed greater variability.

Studies by Lampela (1966), Rouse (1967), and Smith (1964) confirmed the conclusion that predictors of elementary school mathematics teaching success will not be found among the obvious variables of experience, academic preparation, and mathematical knowledge. It might be that the mathematical knowledge of elementary teachers is uniformly so low that it does not play an important role in effectiveness. This should be watched as the academic preparation of teachers improves under the influence of new teacher education proposals.

A more sophisticated approach to the identification of effective elementary school teachers was applied by Peskin (1964). On the basis of standardized tests, seventh grade teachers' attitudes and mathematics understanding were classified as high, middle, or low. A similar appraisal of student achievement and attitudes was followed by investigation of the interactions among teacher and student attributes. This analysis produced several interesting and significant correlations. For example, best student achievement in geometry occurred in classes of teachers with middle range attitudes toward mathematics and high mathematical understanding. Poorest student achievement in arithmetic occurred in classes of teachers with low attitudes and high understanding. This type of knowledge-attitude-achievement interaction seems to be a promising source of deeper insight into the teacher characteristic problem.

Turner (1964) reported progress in a different kind of search for

complex interactions of teacher variables. He analyzed teaching as consisting of three types of tasks: 1) the setting of tasks for pupils, 2) appraisal of pupil responses according to criterial responses, and 3) instructional tasks designed to close the gap between observed and expected student performance. Next he advised Mathematics Teaching Tasks (MTT), a paper-and-pencil test of teacher problem-solving ability that was shown to differentiate between experienced and inexperienced teachers. In the main study, involving beginning teachers in 12 Indiana school districts, Turner sought significant interactions among scores on the MTT, teaching success (measured by supervisor ratings), and various socioeconomic indices (measured by Ryan's Teacher Characteristic Schedule) concerning the teacher and the institutional context within which he worked. Although none of the findings were considered conclusive, the study suggested a variety of hypotheses about the way teacher success and MTT ability change during the early years of teaching and the ways these factors are influenced by the socioeconomic makeup of the community in which the teacher works.

The most obvious conclusion of the investigation was that the question of predicting teacher effectiveness is not simply answered by direct measurement of obvious variables, but must be viewed as a complex interaction of several interrelated classes of variables. This is a good observation for guiding those interested in future teacher characteristic and effectiveness research.

### *Teacher Classroom Activity*

The classroom behavior of mathematics teachers has been a popular topic of research for many years. However, the results of comparing various name methods of teaching—such as discovery, laboratory, lecture, etc.—are generally inconclusive and unreliable.

#### *Methods of Instruction*

"Teaching by discovery" is the catchphrase of current mathematics instruction. As a result, discovery-traditional comparison was the most popular type of methods research in the 1964-1968 period. These studies are examined in another chapter of this *Review*. However, the central question in criticism of discovery method research is applicable to research in other "name" methods: Do the popular method names refer to clearly defined or reliably repeatable patterns of classroom instructional activity?

The emerging viewpoint is that any natural teaching strategy consists of a variety of instructional type components. The proper direction of research, then, is to determine the differential impact of each type rather than to seek some single method best for all situations. Thus, methods

research is shifting toward investigation of the complex interaction between content, teacher behavior, student aptitudes, and student outcomes. The objective is to find the specific teacher behavior style most effective for specific mathematical topics and specific kinds of student population. This emphasis, part of growing interest in individualization of instruction, is exemplified in studies of Aptitude-Treatment Interactions (ATI).

*Aptitude-Treatment Interactions.* The first challenge in ATI research, still in an exploratory stage, is to show that instructional treatments can be constructed that give evidence of ATI on relevant outcome measures. Kropp, Nelson, and King (1967) devised four packages of instructional materials for teaching elementary set concepts: 1) verbal-deductive, 2) verbal-inductive, 3) figural-deductive, and 4) figural-inductive. The material was presented to 400 elementary school students in two days. A battery of aptitude tests and a 24-item criterion test (containing equal numbers of verbal and figural items) were administered to test for aptitude-treatment interaction.

Test data showed that for heterogeneous groups the treatments were equally effective; there was an interaction of inductive-deductive teaching treatments with inductive-deductive reasoning ability significant at the .05 level. That effectiveness of verbal and figural treatments was not noticeably related to special student aptitudes might be explained by inadequate distinction in the treatments, invalid aptitude measures, or an irrelevant verbal-figural dichotomy in the chosen learning task.

Behr (1967) examined the verbal-figural treatment dyad in other content areas with subjects at different age levels. He taught a one-day programed lesson in mod 7 arithmetic to 229 students in an elementary college mathematics course. Half the programs embodied a semantic-symbolic mode and half a figural-symbolic mode. The major result was high correlation between semantic ability and success on the semantic treatment.

Becker (1967) designed two different kinds of instructional treatment for teaching summing of number series. In Treatment A, the learner was given the correct formula, in verbal and symbolic form, for summing a particular series and then given an explanation of the structural relationship between formula and series. In Treatment B, the learning task was broken into many steps leading to discovery of a formula. The treatments were written into programs and administered to Algebra I students, matched in pairs according to high/low verbal and mathematical aptitude. On tests of ability to recall terms, symbols, and formulas, ability to find the sum of  $n$  terms of a series, and ability to devise a formula for the  $n$ th term of a series, there was no evidence of significant aptitude-treatment interaction. As in the Kropp, Nelson, and King (1967) study, lack of interaction effect

might be attributable to irrelevance of the treatment variables for the particular learning tasks.

The importance of ATI effects can be confirmed or denied only after broader exploration of possible treatment and aptitude variables. Use of programed instruction has limitations as a research simulation of teaching methods, but the value of clearly defined, repeatable experiments seems to outweigh these limitations right now. The brevity of teaching programs, although necessary for experimental purposes, casts doubt on the application of findings to long-term teaching methodology.

*Additional References:* Carry (1968); Davis (1967); Ebeid (1964); Moody (1968).

*Psychological Theory and Teaching Theory.* The recent reconstruction of the school mathematics curriculum also led to a vigorous revival of interest in the psychological aspects of mathematics learning, a revival led by Piaget, Bruner, Dienes, Suppes, and Gagné. Such a combination of curriculum and psychological research emphasis seems to reflect a belief that having determined the important mathematical ideas and how children learn them, effective teaching procedure is a routine corollary. There is, however, a growing awareness that the matter is not so simple; translation of learning theory into meaningful guidelines for teaching is an important problem worthy of the best efforts of ingenious teachers. Evaluation of pedagogical theories derived from learning theories is an important, emerging area of research.

In two studies testing the pedagogical implications of Piagetian theory, Toney (1968) and Trueblood (1967) compared teaching methods in which students manipulated concrete instructional materials with methods in which the teacher simply demonstrated or described the manipulation. Results did not significantly support the student manipulation method (presumed to embody the proper psychological principles). However, the Toney study involved a very small sample, and in both studies Piaget's theory was oversimplified and superficially interpreted in the design of the instructional treatments.

Woodward (1966) conducted an interesting experiment to test Ausubel's idea that meaningful verbal learning can occur only when more inclusive relevant concepts exist and are readily available in the cognitive structure of the learner. The idea implies, among other things, that presentation of an advance organizing concept prior to a verbal learning task will enhance learning. Woodward examined this hypothesis and a kind of converse one which proposed that discovery learning would be more successful if overt postorganizers were presented. He taught mod 11 arithmetic to a total of 44 college subjects using four methods—1) preorganizers



and discovery instruction, 2) preorganizers and verbal instruction, 3) postorganizers and discovery instruction, and 4) postorganizers and verbal instruction—each mediated by computer. A learning test (administered one day after instruction) and a transfer test (administered one week later) yielded no significant differences or interactions to support any of the hypotheses. An interesting next step in this direction would be comparison of those results with yet a fifth treatment, no pre- or postorganizers, reflecting Hendrix's admonition about premature verbalization and its effect on discovery learning.

Proctor (1967) tested the hypothesis that student learning and class participation would increase significantly if 1) specific objectives were clarified for the student before presentation of the learning task, 2) effective feedback apparatus was available during the learning task, and 3) there was an objective relation between achievement of specified behavioral goals and course grades. Proctor derived these hypotheses from Gagné's theory that every learning task can be described in terms of desired behavioral objectives and analyzed into a number of prerequisite, simpler objectives. His experimentation confirmed the learning aspect of the hypothesis, but the level of student participation in class was not significantly affected by implementation of the stated procedures.

### *Media of Instruction*

Programed instruction plays an important research role as a means of stimulating classroom instructional methods under rigorously controlled, repeatable conditions. However, programed instruction was originally conceived as a replacement for traditional classroom teaching—as a tool that could offer pacing and feedback on an individual basis. Since mathematics has been a popular topic for developers of programed material, there has been a great deal of research on the effectiveness of this new instructional medium. The most frequently debated and tested question was about the relative merits of programed and teacher-directed instruction. Results of media comparison studies have failed to establish the superiority of programed or conventional instruction. In fact, the evidence is sharply contradictory. Zoll (1969) reviewed research in programed instruction and found that in 13 comparative studies differences favored programed methods in 3 cases, traditional methods in 3 cases, and neither method in 7 cases.

Research in programed instruction has, quite properly, moved toward defining the interactive effects among various types of programing, subject matter, and learner characteristics and toward determining how programed materials can be most effectively used in conjunction with standard teaching procedures. For example, Morgan (1965) and Callister (1965) examined stress, anxiety, and achievement interactions in programed and

conventional classes; Wiebe (1966) tried to find the most effective combination of programmed and teacher-directed instruction with low-achieving students. It seems fair to conclude that, far from supplanting teacher classroom instruction, programming will probably prove to be an important instructional device to be effectively integrated into a total instructional plan.

Television is another instructional medium that seemed to offer great promise of improving instruction. However, the research evidence on methods and effectiveness of televised instruction in mathematics is slim, particularly in elementary and secondary schools. The best known use of television in elementary school mathematics instruction is Patterns in Arithmetic (PIA), developed at the University of Wisconsin. PIA utilizes television lessons and coordinated teacher manuals and pupil exercise books in an arithmetic program for grades 1-6. The course is now being used by over 135,000 students in eight states.

During the 1966-1967 school year, the Wisconsin Research and Development Center for Cognitive Learning directed an evaluation of PIA in first and third grade classes of selected Wisconsin and Alabama Schools (Braswell and Romberg, 1969). The results were generally favorable to the televised program, and opinion inventories showed that teachers and students liked the televised course.

Nazarian (1967) observed that at the college level most studies have shown television to be as effective as standard instructional techniques—at least on gross measures of achievement. He investigated the response to television of various ability groups. Using two classes each of high and low ability (measured by SAT score) college general mathematics students and an experimental instructional pattern of two television classes and one recitation per week, he found that the control lecture method produced better achievement. High aptitude students did better in the experimental course, but low aptitude students did better in the control course. None of these differences was significant, however.

Evidence or lack of it from this and several similar studies shows that the possibilities of television as a component in mathematics instruction have barely been tapped.

*Additional Reference:* Lane (1964).

### *Teacher-Pupil Interaction*

If teaching methods research is ever to produce reliable results in classroom experimentation, it must develop rigorous behavioral definitions of teaching styles and observational techniques that dependably and accurately measure conformity to the styles in question. Throughout educational research there is a growing interest in systematic, empirical study

of teacher-pupil behavior that shows promise of providing these necessary behavioral concepts and observational methods. Application of these techniques to the study of mathematics teaching in particular has been limited, but promising.

*Interaction Analysis.* Flanders's system of Interaction Analysis is a procedure for quantifying direct and indirect teacher influence in the classroom; Flanders's system has been used to test a variety of conjectures concerning the relationship between classroom climate and student achievement or attitudes. In a study of 16 mathematics classes, Flanders (1965) examined the following hypotheses: 1) indirect teacher influence increases student learning when a student's perception of the goal is confused and ambiguous; 2) direct teacher influence increases learning when a student's perception of the goal is clear and acceptable; and 3) direct teacher influence restricts learning when a student's perception of the goal is ambiguous. All three hypotheses were supported. Successful teachers consistently exerted more indirect influence than direct influence. They also showed a tendency to move from extreme indirect influence at the beginning of a unit study, when goals were more ambiguous, to more direct influence as goals of the unit became clearer.

Interaction Analysis was also used in the Minnesota Study of Characteristics of Teachers That Affect Student Learning (Torrance and Parent, 1966), and it became part of an expanded instrument used to study pupil involvement and mathematical content in the "Five State Project" at the Minnesota National Laboratory.

The "Five State Project" was begun in 1961 as a field study of effectiveness of four experimental secondary school mathematics programs. In a report of one phase, Wright (1967) attempted to describe the impact of experimental curricula on patterns of teacher-pupil verbal interaction. Flanders devised a modification of his original system to determine involvement behaviors, the frequency and roles in which students are drawn into various aspects of classroom activity. A content oriented observational system derived from one of Wright and Proctor's (1961) yielded data about the pattern of content behaviors—the relative emphasis of theoretical and conceptual mathematical activity in experimental and control classes.

Sixty-two observers used the involvement part of the observational system during the spring of 1964 and 1965. The accumulated data showed that pupils in experimental and control classes played nearly the same role—passive or limited to high controlled responses amounting to about 20% of all behaviors. The two significant differences showed that experimental teachers asked more confronting, seeking questions and their students responded more independently. However, this behavior constituted only 4% of all activity.

Twenty-five other observers used the content part of the observational system during 1964 and 1965. The data presented no clear pattern of frequency or sequencing among the content behaviors in either the experimental or control classes. There was slight indication that class discussion in experimental programs paid more attention to theoretical matters than did class discussion in control classes. The discussion in the experimental group did not de-emphasize basic problem-solving skills, however.

The lack of evidence of strong behavioral pattern differences between experimental and control classes might indicate that curricula alone do not exert the hoped for influence on classroom procedure. However, there were four quite different curricula lumped together in the experimental classes. Furthermore, the content analysis system does not seem sensitive enough to record fully and accurately the structure of mathematical discourse.

*Other Classroom Observation Schemes.* Pate (1966) devised a different kind of observational instrument and used it to gather empirical evidence "Transactional Pattern Differences" between classes studying SMSG and more traditional curricula. His results corroborate findings in the involvement aspect of the Five State Project; that is, the predominant pattern of interaction was one in which recall and recognition questions were directed at individual students in a climate of control. Although there was some evidence that SMSG teachers used more divergent questions and analysis questions to elicit spontaneous and creative response, this behavior was only a small part of classroom activity.

The Wright and Pate studies give some indication that suitably designed experimental curricula influence patterns of classroom interaction. Others have begun refining the observational instruments to make them more sensitive to the important factors in classroom discourse about mathematics.

*Additional References:* Buck (1967); Fey (in press); Stilwell (1967).

### *Teacher Organizational Responsibilities*

A teacher's central responsibility is active leadership of classroom learning activities. Thus, research most often focuses on patterns and media of classroom instructional behavior. However, in preparation for teaching and during a given class session, the teacher must make a variety of procedural decisions involving grouping of students, class size, assignments, pacing, enrichment, etc. One particularly important study offered insight into the most effective choices in decisions about ability grouping.

As part of a large-scale study (Goldberg, Passow, and Justman, 1966), Neill (1966) compared the effects of enrichment and acceleration on



achievement of 1477 academically talented junior high school students. Using the Lorge-Thorndike Verbal Intelligence Test and STEP reading and mathematics tests to determine initial status and three different criterion tests, he found that a contemporary mathematics program under acceleration leads to greater achievement gains than an accelerated traditional program or an enriched program of either traditional or contemporary content.

Neill also checked for interaction between various teacher and student attributes and achievement. He found that teacher characteristics contributed substantially less to differences in achievement than student attributes, the main teacher predictor being length of academic preparation. Student attributes closely associated with achievement were intelligence, initial reading and mathematics ability, socioeconomic status, attitudes toward mathematics, and self-appraisal. There was no significant interaction of pupil and teacher sex, but classes taught by men were more successful (a result confirmed in the International Study of Mathematics Achievement).

*Additional References:* Stevenson (1966); Willcutt (1967).

### Summary

Recent research in mathematics teaching has produced no major breakthrough in the search for personal characteristics, education, or classroom behavior of effective teachers. However, several promising trends are emerging in the focus and techniques of research.

First, there is a growing realization that effective teaching is the result of a complex interaction between teacher ability and attitudes and behavior, student aptitudes and attitudes, and the structure inherent in mathematical topics. The traditional search for a simple profile of a composite "good teacher" is giving way to investigations that inquire about what kinds of teaching style and subject-matter organization are most effective for teaching a particular topic to some particular type of student.

Second, there is a realization that student achievement on some standardized test is a grossly inadequate measure for teaching success. More comprehensive diagnostic assessments of student outcomes must be developed and used.

Third, several previously uninvestigated classes of teaching variables have been discovered by exploratory studies. The creative thinking ability of teachers has been suggested as an important determinant of success. Classroom observational techniques offer powerful new methods in the study of teacher-pupil interaction variables. The most effective instructional use of television, programed materials, and computers is as yet

undetermined. The pedagogical implications of recent developments in psychology are just beginning to be explored.

Fourth, the wide range of research results bearing on the activity of classroom teaching must be integrated by a theory of mathematics teaching and translated into programs of teacher education. Although these problems are receiving attention elsewhere in education, little has been done to consider the particular implications for mathematics. Results of research on mathematics teaching have so often been unreliable and inconclusive, but there is promise of more rigorous and creative work in the near future.

### Bibliography

- Becker, Jerry P. *An Attempt to Design Instructional Techniques in Mathematics to Accommodate Different Patterns of Mental Ability*. Doctor's thesis. Palo Alto: Stanford Univ., 1967. Abstract: *Dissertation Abstracts* 28: 957; No. 3A, 1967.
- Behr, Merlyn J. *A Study of Interactions Between "Structure-of-Intellect" Factors and Two Methods of Presenting Concepts of Modulus Seven Arithmetic*. Doctor's thesis. Tallahassee: Fla. State Univ., 1967.
- Braswell, James S. and Romberg, Thomas A. *Objectives of Patterns in Arithmetic and Evaluation of the Telecourse for Grades 1 and 3*. Technical Report No. 67. Madison: Wis. Research and Development Center for Cognitive Learning, Univ. of Wis., 1969.
- Callister, Sheldon L. *Stress, Anxiety, and Achievement Relationships in Programmed and Conventional Algebra and Geometry Classes*. Doctor's thesis. Provo, Utah: Brigham Young Univ., 1965. Abstract: *Dissertation Abstracts* 26: 5860; No. 10, 1966.
- Carroll, Edward M. *Competencies in Mathematics of Certain Prospective Elementary School Teachers*. Doctor's thesis. New York: Teachers College, Columbia Univ., 1964. Abstract: *Dissertation Abstracts* 25: 5134; No. 9, 1965.
- Flanders, Ned A. *Teacher Influence, Pupil Attitudes, and Achievement*. U. S. Dept. of Health, Education, and Welfare, Office of Education, Cooperative Research Monograph No. 12. Washington, D. C.: Superintendent of Documents, U. S. Government Printing Office, 1965.
- Gage, Nathaniel (editor). *Handbook of Research on Teaching*. Chicago: Rand McNally, 1963.
- Goldberg, Miriam L.; Passow, A. Harry; and Justman, Joseph. *The Effects of Ability Grouping*. New York: Teachers College Press, 1966.
- Henderson, Kenneth B. *Research on Teaching Secondary School Mathematics. Handbook of Research on Teaching*. (Edited by Nathaniel Gage.) Chicago: Rand McNally, 1963. Pp. 1007-1030.
- Kennedy, Joseph Wilson. *The Development of a Test of Skill in Solving Mathematics Teaching Problems*. Doctor's thesis. Bloomington: Ind. Univ., 1963. Abstract: *Dissertation Abstracts* 24: 3202; No. 8, 1964.
- Kropp, R. P.; Nelson, W. H.; and King, F. J. *Identification and Definition of Subject-Matter Content Variables Related to Human Aptitudes*. Cooperative Research Project No. 2117. Tallahassee: Fla. State Univ., 1967.
- Lampela, Roland H. *An Investigation of the Relationship Between Teacher Understanding and Change in Pupil Understanding of Selected Concepts in Elementary School Mathematics*. Doctor's thesis. Los Angeles: Univ. of Calif., 1966. Abstract: *Dissertation Abstracts* 27: 1548; No. 6A, 1966.
- Lyng, Merwin J. *Factors Relating to a Teacher's Knowledge of Contemporary Mathematics*. *Mathematics Teacher* 57: 395-97; 1968.
- Massie, Ronald O. *The Construction and Use of a Test to Evaluate Teacher Preparation in Modern Mathematics*. Doctor's thesis. Lincoln: Univ. of Neb. Teachers College, 1967. Abstract: *Dissertation Abstracts* 28: 4027A; No. 10A, 1968.

- Moore, Robert Ezra. *The Mathematical Understanding of the Elementary School Teacher as Related to Pupil Achievement in Intermediate-Grade Arithmetic*. Doctor's thesis. Palo Alto: Stanford Univ., 1965. Abstract: *Dissertation Abstracts* 26: 213; No. 1, 1965.
- Morgan, James H. *The Relationship Between Levels of Anxiety and Mathematics Achievement in Programmed and Teacher-Directed Instruction*. Doctor's thesis. Storrs: Univ. of Conn., 1965. Abstract: *Dissertation Abstracts* 26: 3760; No. 7, 1966.
- Nazarian, John. *A Comparative Study of the Achievement of College Students of Different Levels of Mathematical Aptitude Using Closed-Circuit Television and a Lecture-Recitation Method of Instruction*. Doctor's thesis. New York: N. Y. Univ., 1967. Abstract: *Dissertation Abstracts* 28: 1215; No. 4, 1967.
- Neill, Robert D. *The Effects of Selected Teacher Variables on the Mathematics Achievement of Academically Talented Junior High School Pupils*. Doctor's thesis. New York: Teachers College, Columbia Univ., 1966. Abstract: *Dissertation Abstracts* 27: 997; No. 4A, 1966.
- Pate, Robert T. *Transactional Pattern Differences Between School Mathematics Programs*. *Arithmetic Teacher* 13: 21-25; 1966.
- Peskin, Anne S. *Teacher Understanding and Attitude and Student Achievement in Seventh Grade Mathematics*. Doctor's thesis. New York: N. Y. Univ., 1964. Abstract: *Dissertation Abstracts* 26: 3983, No. 7, 1966.
- Proctor, Charles M. *An Experimental Study of the Relationship Between Certain Theoretically Postulated Elements in Classroom Learning and Student Achievement, Grade Distributions, and the Incidence of Certain Classroom Activities*. Doctor's thesis. College Park: Univ. of Md., 1967. Abstract: *Dissertation Abstracts* 28: 4546; No. 11A, 1968.
- Rouse, William M. *A Study of the Correlation Between the Academic Preparation of Teachers of Mathematics and the Mathematics Achievement of Their Students in Kindergarten through Grade Eight*. Doctor's thesis. East Lansing: Mich. State Univ., 1967. Abstract: *Dissertation Abstracts* 28: 4031; No. 10A, 1968.
- Smith, Wallace R. *The Achievement of Eighth-Grade Students in Arithmetic with Respect to Selected Patterns of Teacher Preparation*. Doctor's thesis. Norman: Univ. of Okla., 1964. Abstract: *Dissertation Abstracts* 25: 3947; No. 7, 1965.
- Toney, JoAnne S. *The Effectiveness of Individual Manipulation of Instructional Materials as Compared to a Teacher Demonstration in Developing Understanding in Mathematics*. Doctor's thesis. Bloomington: Ind. Univ., 1968. Abstract: *Dissertation Abstracts* 29: 1831; No. 6A, 1968.
- Torrance, E. Paul and Parent, Elaine. *Characteristics of Mathematics Teachers That Affect Student Learning*. Cooperative Research Project No. 1020. Minneapolis: Minn. School Mathematics and Science Center, Institute of Technology, Univ. of Minn., 1966.
- Trueblood, Cecil R. *A Comparison of Two Techniques for Using Visual-Tactual Devices to Teach Exponents and Non-Decimal Bases in Elementary School Mathematics*. Doctor's thesis. University Park: Penn. State Univ., 1967. Abstract: *Dissertation Abstracts* 29: 190; No. 1A, 1968.
- Turner, Richard L. *Problem Solving Proficiency Among Elementary School Teachers*. Cooperative Research Project No. 1262. Bloomington: Institute of Educational Research, Ind. Univ., 1964.
- Wiebe, Arthur J. *The Comparative Effects of Three Methods of Utilizing Programmed Mathematical Materials with Low-Achievers*. Doctor's thesis. Palo Alto: Stanford Univ., 1966. Abstract: *Dissertation Abstracts* 27: 1002; No. 4A, 1966.
- Woodward, Ernest L. *A Comparative Study of Teaching Strategies Involving Advance Organizers and Post-Organizers and Discovery and Non-Discovery Techniques Where the Instruction is Mediated by Computer*. Doctor's thesis. Tallahassee: Fla. State Univ., 1966. Abstract: *Dissertation Abstracts* 27: 3787; No. 11A, 1967.
- Wright, E. Muriel J. *Teacher Pupil Interaction in the Mathematics Classroom, A Sub-Project Report of the Secondary Mathematics Evaluation Project*. Technical Report 76-5. St. Paul: Minn. National Lab., Minn. State Dept. of Education, 1967.
- Wright, E. Muriel J. and Proctor, Virginia H. *Systematic Observation of Verbal Inter-*

- action as a Method of Comparing Mathematics Lessons. Cooperative Research Project No. 816. St. Louis: Washington Univ., 1961.
- Zoll, Edward J. Research in Programmed Instruction in Mathematics. *Mathematics Teacher* 62: 103-110; 1969.

### Additional References

- Buck, Crayton L. *Mathematics Teaching Behavior of Selected Intermediate Grade Teachers Utilizing the OScaR (EM) for Systematic Observation*. Doctor's thesis. University Park: Penn. State Univ., 1967. Abstract: *Dissertation Abstracts* 28: 3525; No. 9A, 1968.
- Carry, LaRoy Ray. *Interaction of Visualization and General Reasoning Abilities with Instructional Treatment in Algebra*. Doctor's thesis. Palo Alto: Stanford Univ., 1968. Abstract: *Dissertation Abstracts* 29: 475; No. 2A, 1968.
- Davis, John. *An Investigation of the Interactions of Certain Instructional Strategies with the Structure of Basic Mental Abilities in the Learning of Some Mathematical Operations*. Doctor's thesis. Tallahassee: Fla. State Univ., 1967.
- Ebeld, William T. *An Experimental Study of the Scheduled Classroom Use of Student Self-Selected Materials in Teaching Junior High School Mathematics*. Doctor's thesis. Ann Arbor: Univ. of Mich., 1964. Abstract: *Dissertation Abstracts* 25: 3427; No. 6, 1964.
- Fey, James T. *Patterns of Verbal Communication in Mathematics Classes*. New York: Teachers College Press, In press.
- Lane, Bennie R. An Experiment with Programmed Instruction as a Supplement to Teaching College Mathematics by Closed-Circuit Television. *Mathematics Teacher* 61: 695-97; 1968.
- Stevenson, Robert L. *The Achievement Gains in Mathematics of Seventh Grade Pupils When Achievement Grouping and Flexible Scheduling Are Employed in a Team-Teaching Program*. Doctor's thesis. New York: N. Y. Univ., 1966. Abstract: *Dissertation Abstracts* 27: 3785; No. 11A, 1967.
- Stilwell, Merle E. *The Development and Analysis of a Category System for Systematic Observation of Teacher-Pupil Interaction During Geometry Problem Solving Activity*. Doctor's thesis. Ithaca: Cornell Univ., 1967. Abstract: *Dissertation Abstracts* 28: 3083; No. 8A, 1968.
- Willcut, Robert E. *Ability Grouping by Content Subject Areas in Junior High School Mathematics*. Doctor's thesis. Bloomington: Ind. Univ., 1967. Abstract: *Dissertation Abstracts* 28: 2152; No. 6A, 1967.

### AUTHOR

- JAMES T. FEY Address: University of Maryland, College Park, Maryland Title: Assistant Professor of Mathematics and Secondary Education Age: 29 Degrees: B.S. and M.S., Univ. of Wisconsin; Ph.D., Columbia Univ. Specialization: Mathematics Curriculum Development; Teacher-Pupil Classroom Interaction.





## FOREWORD

Methods and techniques are the ways of doing the work of science. The task of methodology is to appraise generative ideas, methods and products and values by describing and explaining them, by identifying their limitations and advantages, by clarifying underlying presuppositions, and by projecting judgments about the likely consequences of their use in the context of educational research.

This conception of research methodology served as a guide to the authors of this issue. The authors were advised not to provide a comprehensive bibliographic listing because this task is being performed by others (e.g., Educational Resources Information Center). Rather, the authors were expected (a) to describe some of the significant ways people are conceptualizing the phenomena of interest and doing research relevant to education in a specific field and (b) to provide critical appraisals of these ways. The authors were asked to select those ways of thinking about or doing educational research that stimulated either their creative flair (appeared promising to them) or their critical bile (seemed to be overrated or misleading).

Unlike most fields of research, a tradition of disciplined criticism seems largely missing in the field of educational research. Future issues of the *Review* will publish unsolicited critical reviews, and we hope this issue will stimulate other writers in these various disciplines to take on the challenge and extend the arguments about what good research really is like.

The mark of contemporary inquiry in education is increased diversity of research paradigms. Although we approached the fact of diversity by seeking authors in various academic fields, we were aware of areas of interest not yet encompassed in the traditional academic departments (e.g., future policy research—taking the future as an object of study, architecture—taking the design of space as an educational variable). We are also aware that one writer can not encompass the field in which he specializes. Who speaks for all of Psychology? Philosophy? Economics? No one. Each chapter is meant to present a thoughtful selection of work, not a representative sample.

D. Bob Gowin  
Jason Millman  
Issue Editors



# I: RESEARCH METHODOLOGY —A POINT OF VIEW

D. BOB GOWIN  
JASON MILLMAN  
*Cornell University*

In the practical sense of doing, research is a matter of techniques. An event occurs or is made to happen by the researcher and some record of the event occurs or is made. Careful study of the record generates factual statements serviceable as evidence for inferences leading to generalizations, explanations, interpretations, predictions and decisions. Research technique is the process of converting events into records and records into factual statements (including tables, charts and other ways of showing relationships).

In the theoretical sense of reasoning, research is a matter of using concepts, conceptual systems, constructs, models, theories, etc. An event can be studied effectively if the researcher generates a *telling question*. Such questions require key concepts or generative ideas which lead and guide the inquiry. Facts without concepts are blind and concepts without facts are empty (apologies to Kant). Research brings together conceptual systems and techniques. Appraisal of research is methodology (Kaplan, 1964).

Gage (1964) identified three aspects of the research paradigm: the *substantive*, the *methodological*, and the *logistical*.<sup>\*</sup> Substantive means the basic concerns of the research; this category is further analyzed by Gage into concepts, variables, and the ways in which phenomena are selected, abstracted, and focused by the research workers. We call this area the *context for inquiry*. Expanding Gage's category permits a more complete consideration of what must necessarily be thought about within the research paradigm.

Gage's second category is the methodological, in which he included ways of measuring variables and of collecting and analyzing data. We prefer to call this category *methods of work* because, as noted in the Foreword to this volume, we reserve the term methodology for the study of methods.

Logistical, Gage's third category, means the host of concerns with men and money, organization and administration, etc., that are necessary in planning, executing, interpreting, communicating, applying, and improving research. We do not discuss this important category here.

<sup>\*</sup> See Gowin and Strzepek (1969) for a criticism of Gage's article and for a further development of some of the ideas in this present chapter.



Two other components of the research paradigm appear to be important. One of these we call *products*, what the research produces. The outcomes of research are usually thought to be limited to published discussions of the research results. By using the broader category, products, a larger number of aspects of research outcomes can be analyzed. The fourth component is *values*. Explicit treatment of the many ways that consideration of value enters into research is especially important in a field as saturated with values as education is. The tug between values and knowledge-claims is particularly fruitful for analysis. We believe that consideration of research within this larger view of methodology is necessary for significant improvement in educational research.

### *Context for Inquiry*

The *context for inquiry* includes the scene and phenomena of interest, telling questions and principles of evidence, key concepts and conceptual systems, basic assumptions and presuppositions.

By the scene we mean that most inclusive frame of reference in which the phenomena of interest can stand out clearly; the scene is the background for the foreground phenomena. Any research is an abstraction and selection from an infinite variety of possible things one might study. For example, sociologists Harp and Richer identify social interaction as the most inclusive sociological domain (Chapter 8, this issue). When they turn to educational research they note a quality which colors the scene: "The pervasive character of an ideology of equal opportunity and the continuing role assigned to education as an avenue of mobility in American society could not help but influence the direction and character of sociological research in this area."

It is our belief that the analysis of the scene in any research area will reveal some pervasive quality which helps determine the direction and import of research, and that the identification of this pervasive quality is useful in assessing the logic of the moves back and forth between scene and knowledge (Dewey, 1930).

Emerging from the background scene are the foreground phenomena of interest. The phenomena of interest are simply the more specific researchable domains within a larger scene. In the case of sociology of education they are identified by Harp and Richer (Chapter 8, this issue) as social organizations, collectivities and institutions.

Phenomena of interest are approached in two different ways, ways which appear to correspond to the origin of the inquiry. When "enquiry has its origin in conceptual structure . . . we are able to formulate a telling question" (Schwab, 1962, p. 12). A fool can ask more questions than seven wise men can answer, but a *telling question* is one which "tells" on the phenomena and facilitates inquiry. Telling questions help convert what is

puzzling about a situation into something that can be thought about intelligently. There is a close relationship between the telling questions and the key concepts, generative ideas, and conceptual systems. Growing out of the scene of American race relationships is a phenomenon of interest, Negro education of the south. (See Chapter 2, this issue, Beach.) The conceptual model of "conflict—unintention—accommodation" relates to the telling question: In history, are the unintended consequences of man's willful actions as significant as the changes resulting from the "force of intention"?

Inquiry may also have its origin, as John Dewey believed, in problematic situations. Still other writers suggest that inquiry may be brought about by a puzzling observation, by access to a new source of data, subjects, instruments or techniques of study, or by existence of "previously unexamined phenomena" (Chapter 2, this issue, Beach).

Most of the phenomena of interest in educational research appear to originate from such ordinary situations and to be approached with questions that are not very telling. An example of such a question is, "What are the faculty and student subcultures and their associated norms?" (Chapter 4, this issue, Sindell.)

Once a question, telling or not, is formulated, the generalized phenomena of interest can be converted into a specific aspect of that phenomenon which will count as evidence. Any question remains ambiguous until the sort of thing which would count as an answer is agreed upon.

There is a crucial connection between concepts and facts. Facts may be given three distinct but related meanings. Fact One is an event which occurs (either being made to occur by the researcher or merely happening). This event must leave a record in order to be chosen for study. Researchers can spend much time profitably thinking about techniques for making records of events. The record is Fact Two. A good device for recording events generates an index to the phenomena of interest. Fact Three is a factual statement, typically in verbal or mathematical form. Factual statements are based on records of events occurring in the phenomena of interest. Since concepts can be related to facts in all three senses (the events, the record, and the statement), it is important for purposes of test, analysis, and criticism to be clear about which level of fact one is dealing with.

One may commit a Type I or II error in relating concepts to data: one may become too specific and lose important evidence; or one may become too loose and include so much that it is impossible to make a legitimate discrimination. Precision and clarity in a conceptual system are commendable, but care must be taken that the legitimate complexity and richness of the phenomena of interest are not lost. The enormous complexity of educational phenomena is a condition of work in education.

How shall *concept* be defined? For all of its frequent use in educational

writing, little has been done to clarify its conflicting meanings. In general there is a tension in meanings between the psychological (studies in concept formation) and the logical (concept as a class, as a construct, as a carrier of meaning). A child possesses the concept of "liquidity" when he correctly identifies milk, oil, honey, water as liquids; he sees what is common in events which are quite different. Thus, there are three elements tied together in the meaning of concept: the stability of response, the linguistic sign, and the commonality in different events. We stipulate the following definition of *concept*: a concept is a sign which points to a commonality in events and which permits the concept user to make relatively stable responses to those varied events. The signs which are vehicles for the concept are largely linguistic and conventional. The commonality in events may range from simple similarities to regularities to law-like invariance.

Conceptual structures are logically organized sets of concepts. Conceptual structures serve many roles depending on the phase of the research. They help the researcher formulate the telling question, identify index phenomena through the principle of evidence, formulate knowledge-claims, and interpret products. To identify a conceptual structure is to understand the power and limitation of inquiries that take place under its guidance (Schwab, 1964).

When these elements in the context for inquiry are analyzed, it is possible to expose the basic assumptions and presuppositions of the inquiry. An assumption is what is taken for granted to establish meaningful communication. What is taken for granted in one study may become an object of inquiry in another study, but the second study will necessarily require another set of assumptions.

When Beach, for example, looks at the context for historical research in education he finds a number of assumptions and presuppositions (Chapter 2, this issue). One of these is the assumption that education is something done to people rather than something they seek and gain on their own. Beach wonders whether historians can escape to some extent the myopia which makes historians both distinctively and unimaginatively professional.

We do not want to argue these issues but rather to point out that any set of knowledge-claims rests upon assumptions and presuppositions. Analysis of these elements is basic philosophical work.

### *Methods of Work*

A *method* is a procedural commitment, a collection of techniques or ways of doing things that may be generalized or made common to a variety of situations. Method involves regular steps, planned sequences, ordered phases, and related stages of inquiry. Research methods are of many kinds: experimental, empirical, historical, pedagogical, comparative, conceptual.

Techniques of work are the many different specific ways of working. A criticism of research methods in a given work involves the logic of the moves from the occurrence of the event to the record of the event, to the specific tasks of inquiry.

Questions of the validity of methods of work raised by the authors of chapters in this issue include: Is the prose of other men which remains to be read representative of either the author or some group with whom he is associated? (Chapter 2, Beach.) Should mentalistic goals be salvaged from behavioristic criticism? (Chapter 5, Waks.) Do participant-observers in the classroom lose objectivity as they become emotionally involved? (Chapter 4, Sindell.) Do the tools of measurement of situational differences seem to be measuring the wrong variables? (Chapter 9, Mitchell.) Is a series of univariate analyses of several dependent criterion variables inappropriate? (Chapter 10, Tatsuoka.).

### *Products*

If we ask about the *products* which research generates, instead of merely about its findings, we have a broader base of analysis. The typical products are conclusions and decisions (Cronbach and Suppes, 1969) or, more specifically, facts, generalizations, concepts, theories, explanations, predictions, and interpretations. In addition, an inquiry may produce a new question, a new area of uncertainty, a puzzling observation, and a new set of terms. A product may be simply propositional, as most works are, or more tangible such as a set of instructional materials. A product is what the investigator has when the work is finished that he did not have when he began.

Complete analysis of inquiry demands as careful scrutiny of the product of the research as of the context for inquiry and methods of work. An analysis of products includes not only identifying the product or products (e.g., whether explanation or description, fact or concept) but also involves applying criteria of excellence appropriate to the several, specific products. If the products are tangible, such criteria include durability, cost, and effectiveness. Consistency, simplicity, and explanatory power are criteria by which theories, for example, are judged.

### *Values*

*Values* are created at all phases of disciplined inquiry. Values constitute both a separate category and an aspect of each of the other three categories. As a separate category the value questions concern the identification of values, the relation between intrinsic and instrumental values, and the identification of within-the-field values and outside-the-field values. As an aspect of each of the other three categories, the value questions emerge from the specific tasks of inquiry.



In general, values are deliberated interests, justified likings. To identify values, expressions of likings, desires, interests, etc. must be located—not always an easy task. When such expressions show conflict, a principle of decision must be invoked.

The most critical value judgment concerns the significance of the entire research effort. "So what?" questions typically have two answers. One claim is that there was *intrinsic value* in performing the research. Doing research does have intrinsic value; satisfying curiosity is curiously satisfying. The other claim is that there was *instrumental value* in the research products. The outcomes of research are thought to be good for something else: reaching a decision, solving a problem, improving a practice, or stimulating further inquiry.

We believe it is very important for workers to search for and examine the intrinsic values in their field. These values are in much need of explicit recognition and justification. The chief reason for becoming explicit about these intrinsic, within-the-field values is that unless the field has some claim to intrinsic value it may become extinct because the instrumental values it serves may be served by other means. When history, for example, is justified only by its instrumental value in creating patriotic citizens, it lays itself open to abandonment when other means are found equal or superior to achieving love of country. When psychology is justified only on grounds that it helps students understand themselves better, it may be replaced by the guidance counselor or psychotherapeutic encounter groups.

Each of the other three categories has value considerations. In the *context for inquiry* perhaps the most fundamental value orientation of the scene and phenomena of interest is "What knowledge is of most worth?" Every chapter in this issue gives one or more answers to this question. Telling questions and conceptual systems have value in the guidance of inquiry. Other recognized values of conceptual structures are these: clarity of meaning, coherence of ideas, accuracy and precision of reference, power of reasoning.

Values found in *methods of work* include the following familiar list: reliability, systematic procedure, repeatability and replication, technical competence, efficiency, economy, and intersubjectivity.

For *products* other values appear. Research outcomes may have commercial value, aesthetic value, moral value, pedagogical value. These are examples of instrumental values. Within the pattern of inquiry, however, the research product is valued which provides a clear explanation of something which has been puzzling; so is a warranted generalization which helps to connect elements in educational experience which had heretofore been seen as separate. A well-argued prediction is valued since it helps decision-making in the present. Perhaps the value of knowing the truth transcends all others.

## *Afterword*

What is the value of our point of view toward educational research?

In the Foreword we noted the increasing diversity in the way educational phenomena are viewed. The point of view outlined here is a way to bring together in some comprehensible fashion these different approaches to research. Many pleas have been made for interdisciplinary research, but few means have been found to communicate accurately and well between fields. A common set of terms and distinctions is needed to facilitate discussion about inquiry in different areas. We think that our point of view is a step in that direction and we hope that others will challenge and change it and contribute toward a more adequate view. We see improvement in three major areas.

First, a point of view will help in the guidance of research practice because it will help to develop thorough-going and intelligent criticism. We have in mind here the careful analysis of both successful and unsuccessful research practice in much the same way as any practice, such as farming or painting, may be improved as a consequence of ferreting out the conditions of success and failure in each particular case.

Second, a point of view will help in the guidance of research policy based upon an analysis of priorities, and what is important and why. This goal is facilitated by the broader framework for criticism of research proposals and research products. If criticism of research proposals were customarily set in the framework suggested here, the merits of proposals would be more validly assessed.

Third, a fully-developed point of view will be a research theory, i.e., a philosophy of educational research. Some philosophical view of science lies behind or is explicit in every research effort. If it is assumed that distinctive phenomena require distinctive approaches, then one must determine what is distinctive about educational phenomena. The distinctions necessary to solve problems in the conduct of physical inquiry may actually blur or distort the distinctions necessary to solve problems in another context, such as educational inquiry. To the extent that this point of view is correct, the need for a special philosophy of educational research is clear. The categories of such a research theory include the context for inquiry, the methods of work, the products, and the values adumbrated in this essay. The source of meaningful content would come from continued critical analysis of completed studies. It is gratifying to us to read the examples of disciplined reflection upon research practice represented by the authors of this volume. More work of this kind would go a long way toward satisfying a critical need for a theory of educational inquiry.

## Bibliography

- Cronbach, Lee J. and Suppes, Patrick (editors). *Research for Tomorrow's Schools: Disciplined Inquiry for Education*. New York: Macmillan, 1969.
- Dewey, John. Qualitative Thought. *On Experience, Nature, and Freedom*. (Edited by R. J. Bernstein.) New York: The Liberal Arts Press, 1960. Chapter 8, pp. 176-99.
- Gage, N. L. This Side of Paradigms: The State of Research on Teaching English. *Research Design and the Teaching of English: Proceedings of the San Francisco Conference, 1963*. Chicago: National Council for Teachers of English, 1964. Pp. 22-31.
- Gowin, D. B. and Strzepek, J. The Far Side of Paradigms: Conditions for Knowledge Making in English Education. *The English Record*. Binghamton: New York State English Council, Fall 1969.
- Kaplan, Abraham. *The Conduct of Inquiry*. San Francisco: Chandler, 1964.
- Schwab, J. J. Structure of the Disciplines: Meanings and Significances. *The Structure of Knowledge and the Curriculum*. (Edited by G. W. Ford and Lawrence Pugno.) Chicago: Rand McNally, 1964. Pp. 1-50.

### AUTHORS

- D. BOB GOWIN Address: Cornell University Title: Associate Professor of Social Foundation Age: 43 Degrees: B.A., Univ. of Tex.; M.A., Stanford Univ.; Ph.D., Yale Univ. Specialization: Philosophy of Education.
- JASON MILLMAN Address: Cornell University Title: Professor, Educational Research Methodology Age: 36 Degrees: B.A., M.A., and Ph.D., Univ. of Mich. Specialization: Educational Measurement; Research Design.

## 2: HISTORY OF EDUCATION

MARK BEACH

*University of Rochester*

Most historians, including those concerned with education, have been reluctant to give careful attention to their methods. They have been especially hesitant to philosophize about their work. That is, historians have not been anxious to seek the assumptions underlying their research, to explore the implications of these assumptions, and to think through their bases for making statements of fact, values, and social utility. When they do talk about methods, they tend to discuss techniques for locating and verifying data and, to a lesser extent, drawing conclusions from these data. Discussions of that nature are, in the parlance of the profession, conversations about historiography.

Cremin's (1965) use of the term *historiography* differed radically from those more conventional meanings. His concern, like Bailyn's (1960), was with the phenomena of interest to the researcher, and how the traditions of historiography of education have unduly restricted our view of those phenomena we consider to have influenced education. Cremin was convinced that education must be taken to mean far more than schooling, that inquiry into the history of schooling must transcend the study of rhetoric, and that the rhetoric of school reformers must not be taken as synonymous with historical truth. In short, Cremin addressed his remarks to the key historiographic questions, indeed, the key questions of any research—those at the beginning. How do researchers select what to study? What cultural, professional, and personal biases do they bring to the questions they choose to investigate? How do those biases influence their definitions and procedures? Can researchers escape to some extent the myopia which makes them both distinctively and unimaginatively professional?

Such questions as these, first in the sequence of research, will be considered first in this essay. After discussing some recent literature from that standpoint, I consider the gathering and analysis of historical data and the use of these data in making factual and theoretical assertions. In my concluding remarks I call attention to trends in the recent literature and suggest lines of further inquiry.



## *Selecting and Defining Topics for Inquiry*

Current social issues remain the most popular sources of problems for historical investigation. Words such as *revolution* and *conflict* symbolize both professional and civic concerns. Two of the most provocative and influential recent volumes addressed to social issues are Jencks and Riesman's (1968) *The Academic Revolution* and Feuer's (1969) *The Conflict of Generations*. Both books deal primarily with higher education; both reflect the essentially sociological cast of mind of their creators; and both were written by historians who lack extensive formal training in the discipline. Such amateurs are sometimes hampered by lack of historical skills and knowledge. That is the case with Jencks and Riesman, whose text is cluttered with historical inaccuracies. Often, however, amateur efforts are marked by a humanism, a creative imagination, or a scholarly technique or insight imported from some other discipline. Feuer's book is an example of the best of this tradition. Unlike Jencks and Riesman, his history is sound and to the point, and his analysis is based upon thoughtfully selected and applied concepts from the social sciences. Other topics of current interest that have recently been explored historically are loyalty oaths (Gardner, 1967), the relationship between bureaucracy and the problems of urban education (Katz, 1968a; Tyack, 1967a), and reactionary responses to radical proposals and actions in portions of the work by Karier (1967). Occasionally there is a book which achieves greater contemporary significance than was originally anticipated. Ambrose's (1966) perceptive history of West Point, written in the context of American attitudes about a professional army, is an example of such a study. Another example is Kevles's (1968) description of the development of intelligence testing during World War I.

The simple desire to acquire knowledge about previously unexamined phenomena has been a second major impulse to recent historical inquiry. There are few events or lives or movements of any consequence which have no chronicle. But there are many whose accounts are totally inadequate. New evidence or techniques of analysis require that these subjects be treated from a fresh perspective. Fisher's (1967) portrayal of the industrial education movement, for example, sprang from her recognition that the movement had not received the attention it deserved. Graham's (1967) history of the Progressive Education Association and Bowers's (1969) record of radical educational reformers during the Depression treat in detail matters previously discussed by Cremin (1961) and others. The outstanding recent example of enlarging and refining previous perceptions is Bragdon's (1967) brilliant study of Woodrow Wilson during his years as a student, professor, and university president. Bragdon's intimate portrait combined prodigious research with subtle inference. It is an example of history at its literary best.

Biographies have been an important subcategory of histories written out of a sustained curiosity combined with a conviction that the subject merits the effort. They have been written from several vantage points. Joncich (1968) described the life and thought of Edward L. Thorndike. She commented extensively upon his work, illuminated the complex ingredients of a successful professional life in the twentieth century, and revealed much of interest in the general history of psychology and education. Tyack (1967b) also portrayed the interplay between man and environment in his biography of George Ticknor. His emphasis was on Ticknor as a personality which both represented and commented upon nineteenth century American life. In Bragdon's (1967) book, Woodrow Wilson occupied center stage simply because of his supreme importance in American history. Other important biographies include Drost's (1967) on the social efficiency expert David Sneddon and Sparrow's (1967) on Mark Pattison and his involvement in the nineteenth century reform of the ancient British universities. There is also a volume that is not biography in the traditional sense, but rather is an intensive account of a small portion of one man's life. Rosenberg (1968) traced in detail the trial of Charles Guiteau, assassin of President James Garfield. Rosenberg's purpose was to illuminate the concepts of human nature, legal responsibility, sanity and insanity, and professional mental care which were the key to determining Guiteau's guilt or innocence. Rosenberg artfully used biography to accent critical questions about the human mind before the birth of scientific psychology in this country. In somewhat the same vein, Rader (1966) used his account of the life of Richard T. Ely to demonstrate the growing alliance between social reform and scholarly expertise during the early years of progressivism.

Histories of specific schools might be termed biographies of institutions. But they usually prove much less satisfactory than accounts of individual lives. This is due largely to the chauvanism which frequently underlies institutional histories. Ambrose's (1966) history of West Point is a happy exception to that general rule, as is Barnard's (1968) portrait of Oberlin College. Callcott's (1966) story of the University of Maryland is also free from the reverence that often characterizes official histories. Finally, Parsons' (1969) picture of Cornell is unusual because he sees the growing personality of the campus in architecture and physical planning. Pembroke and Douglass Colleges and Illinois, Wayne State, St. Cloud, and Bowling Green universities have also been the subjects of recent more traditionally conceived institutional histories.

Most of the studies discussed thus far are of general and professional interest. They are stories of well-known people, events, or institutions. They bring history to bear upon issues which are familiar to the literate public, issues such as the nature of intelligence, the implications of a professional army, and the social functions of higher education. The

items mentioned in the next several paragraphs are more specialized. These studies fall roughly into three categories: those whose authors seek meaning from ideas or events that previously seemed unrelated, those written to synthesize old data or merge it with new information in an effort to provide new intellectual frameworks, and those written as bold challenges to old interpretations and concepts.

Studies in which researchers try to make sense out of ideas or phenomena that had appeared unrelated or unexplainable usually are characterized by the use of some powerful concept or telling question. Consider, for example, Wilson's (1968) attempt to provide common ground for the work of five widely diverse intellectuals: Charles Sanders Peirce, James Mark Baldwin, E. A. Ross, G. Stanley Hall, and Josiah Royce. These men had little contact with each other, yet the work of each was the work of all. Wilson found each man fully aware of the "hopeless insufficiency of the isolated individual" in an impersonal, urban society. They were united in a "quest for community" in the face of Darwinism and industrialization. Smith (1967) was concerned with some of the same questions for a different era in his report on Protestant influence upon public schools and upon American nationality. Bullock's (1967) report of Negro education in the South was based upon an intricate concept about historical trends rather than about a given situation. He was convinced that the force of accident is as powerful in history as the force of intention. The unintended consequences of man's willful actions, therefore, often bring about significant change. In the case of American race relations, Bullock saw a "model of conflict-unintention-accommodation" which both explains the past and provides hope for the future.

The explicit concepts or theses characterizing the studies mentioned in the previous paragraph tend to make them both coherent and a bit artificial. The concepts make clear why certain questions were asked or certain data sought. They also prevent other and perhaps conflicting interpretations from coming to light. This tendency is not so characteristic of writing done to synthesize and make intelligible a mass of previous work. Wishey's (1968) history of the ideology of the child nurture in this country and Wiebe's (1969) suggestions about the historical relationships between schools and American society are good examples of this genre. Wiebe's work especially reveals thoughtful contemplation of the history of American education. It is a fine synthetic effort, because it leads naturally to further significant questions without necessarily prejudging the answers.

The most fundamental questions Wiebe raised concern the role of institutions in American life. These questions are also at the center of a prime example of revisionist history, Katz's (1968b) discussion of early school reform movements. In that book Katz tried to remove the halos historians have given leaders of the common school movement. He depicted those leaders as benevolent despots more concerned with preserving



their positions in a fragmenting society than with building schools to help individual men live satisfactory lives. Thus school reformers have, in Katz's view, used institutions in a conservative rather than a liberating fashion. Further, they have been interested in reform only when the problems of urban life appeared too pressing to be ignored. Katz saw parallels between past and present behaviors, and lamented that mere humans have not somehow managed to institutionalize his vision of humanity. Katz's revisionism lies in his imaginative restructuring and reinterpretation of old data. Another example of a revisionist interpretation is West's (1967) contention that professional schoolmen, acting in their economic self-interest, were the single most important force behind the abolition of rate bills and the passage of compulsory school attendance laws in the nineteenth century.

Revisionism is also practiced by historians with new evidence gathered to shatter old myths. For example, consider Fishlow's (1966b) analysis of the actual increase of school attendance and support during the common school revival. Fishlow's writing lacks the moral fervor of Katz's, but is of equal historiographic consequence. Separating the rhetoric of reform from its results, Fishlow demonstrated that most of the increases in school attendance during the ante-bellum years happened in the rapidly growing West and South rather than the more stable New England. The "revival" for Fishlow was mostly a paper movement whose importance diminishes as one frames different questions concerning its nature.

### *Seeking, Gathering and Summarizing Data*

Most historical writing is based upon the prose of other men. Evidence such as that found in books, articles, reports, diaries, and private letters has the obvious quality of having been preserved and become available for study. But there are hazards in its use. One danger is the assumption that what remains to be read is somehow representative or typical of either the author or some group with whom he associated. Consider the problems Lannie (1968) faced in trying to sort out the political, religious, educational, and personal ingredients of the New York City school crises of the 1840s. How could he test whether a political leader such as Governor Seward believed his own public statements? The standard technique is to check public statements against private ones, and Lannie did this during a thorough examination of Seward's papers. But private papers are seldom available, and if available, often not used. Therefore, as in the case of Wilson (1968), historians frequently check for reliability by seeking consistency of statements over time.

A check for consistency of statements is also the standard procedure for establishing the relationship between individuals, that is, placing them in groups. Yet the way historians group phenomena determines in



large measure the truth of statements about their nature. Very few historiographic problems are as difficult as those imbedded in the establishment of categories. Every man's taxonomy reflects his view of the universe. Categories tend to take on the attributes of facts, despite their roots in assumptions. Bowers (1969), for instance, had a view of radicals in education which allowed him to include such men as Counts, Kilpatrick, Childs and Rugg, but almost to ignore such men as Hutchins, Meiklejohn, and MacLean, all of whom also flourished during the 1930s. That is, Bowers's notion of radicals did not include men whose efforts were focused on college and universities rather than public schools. Another example of the influence of categories upon research is Katz's (1968a) depiction of a struggle for control of education in Boston between professional schoolmasters on one side and non-professional administrators and school committee members on the other. This division was consistent with his effort to link professionalism with stultifying bureaucracy. Had he a different story to tell, he might have divided the same men between urban and rural birth or between college and non-college graduates and found equally attractive evidence for some other thesis.

Daniels (1968) avoided some of the risks of intuitive or deductive categorizing by establishing quantitative criteria. His problem was to define "scientist." He solved it by including in his study only men who accounted for a specified number of publications in the leading journals of their day. Identify the men, read their reports of their research, and you have "Science in Jacksonian America." But the assumptions are still there: assumptions that exclude leading thinkers and include minor functionaries because the latter published and the former did not; assumptions that define "leading journals"; and assumptions that prescribe publications as the means for diffusing ideas.

Several other problems of evidence center on the issue of validity. Most of them fall within one of two domains: Given specified ideas or events, 1) how representative were they of larger groups of people or series of events, and 2) what impact did they have, if any, upon subsequent human thought or behavior? The first is a question of correlation; the second a question of cause.

The problem of the representativeness of ideas is partially a problem of available evidence. Historians are limited to what evidence remains and therefore are tempted to confuse the merely articulate with the truly representative. Thus, it is easy to assume that the man who appeared to be a leader actually had followers, the man who claimed intellectual kinship with others actually had kinsmen, or the man who reported having been influenced actually was. Wishy (1968) and Hayter (1968) chose imaginative ways of dealing with this problem. Wishy checked the models of child nurturing and behavior he found in prescriptive manuals against

the idealized behavior of children and their parents found in popular literature and textbooks. Hayter examined the options and behavior of farmers who had been exposed to agricultural education. Both authors used a measure of consumer receptivity to ideas as one test of the representativeness of those ideas. This technique would have significantly improved several other recent efforts. For instance, both Fisher (1967) and Calvert (1967) discussed education for students who were old enough to record their responses to their experiences. Each study would have profited from a comparison of those student records with the concepts and claims of advocates of practical education. Such a comparison might have revealed two quite distinct views of industrial or technical education: one held by its vocal proponents, the other by those who used it as a route to upward mobility.

The representativeness of events, the extent to which their causes and outcomes correlate with other apparently similar events, is an equally difficult matter to assess. This problem lies at the heart of the oldest and most extensive historiographic debate in the profession: the discussion over the uniqueness versus the generality of historical phenomena. Consider Katz (1968b) and Rosenberg (1968), two recent examples that help put the matter in concrete terms. Both wrote superlative history. Each is on an opposite side of the issue. Katz wrote that the American elite has used public schools as agencies of repression during each major phase of adjustment to industrialization. Two conclusions follow from that premise. One is that the "imposition of reforms" in the 1840s was similar to current efforts to change urban schools. To study that period, then, is to study ourselves in the hope that we will do better by our children than our highminded despotic ancestors did by our fathers. The second conclusion is that one can generalize about the plurality of reforms during a period by examining particular examples in great detail. In other words, history is a guide to present action, and its lessons are adequately learned by the case method.

In contrast to Katz, take Rosenberg's fascinating narrative of nineteenth century concepts of the mind and mental illness. These concepts and their implications are no less significant for the understanding of nineteenth century education than are the motives which Katz imputes to his reformers. The difference is Rosenberg's implicit conviction that the case of Guiteau was unique, and that therefore his book about it could be approached as an art form rather than as a pedagogue with pages. For Rosenberg, but not for Katz, historical events take place in sequence which are relatively discrete. Further, each sequence represents a unique set of circumstances. For both reasons there is probably little utility in the study of history.

Whether ideas and events are unique or representative of a larger

domain is a particularly vexing problem for historians using materials relating to local history. Historians frequently use examples from small areas to illustrate and make more concrete large concepts and conclusions. Smith (1967) is an exemplar of this technique. On other occasions, local history is important for the understanding of some subsequent event of more general interest. Tyack's (1968) sketch of the background of the Pierce Case is an example. The danger is that the illustrations or background information often turns out to be also the total amount of evidence.

The study of local history is one way to seek extensive data from a variety of standpoints and from sources whose relationship to each other can often be traced in detail. In the opposite direction, but toward the same goal of resolving the unique versus general dilemma, is the increasing use of data representing masses of people. Census reports are the most obvious sources of these data for the historian. The use of census data, state financial records, and other sources of large-scale public information is illustrated by Fishlow (1966a and b) and Stambler (1968). Such useful sources, however, bring with them their own historiographic problems. Any one familiar with collecting and compiling numerical data knows of the inaccurate counts, faulty transcriptions, confusing and non-comparable categories, and idiosyncratic statistical procedures which constantly harass the historian willing to undertake this study. Nevertheless, the mere necessity of coping with such problems, frequently with the aid of some machine for data processing or computing, may have a salutary effect upon those venturesome scholars. They may emerge from the experience with questions more sharply framed, data more precisely defined and clearly organized, and analytical procedures more carefully thought through. For instance, Bidwell (1966) established categories and filled them with quantified data in his examination of the roles of social homogeneity and control of schools in moral education. Benson's (1968) article on scientific procedures in the study of past public opinion is an example of approaching historiography itself from the standpoint of experience with quantitative data. For further development of this theme and other applications of computers in historical research, see Clubb and Howard (1967).

The difficulties in trying to correlate phenomena also exist in the attempt to ascertain causal relationships. To show cause, the researcher must show both correlation and at least a necessary relationship between phenomena. Further, he must be careful to distinguish between imputing and assessing causes. Proving that ideas or events are related in such a way that the existence of some can be attributed to the prior existence of others is the most difficult task facing any historian. It is especially rigorous for an historian who believes in the uniqueness rather than in the continuity of historical phenomena. Yet virtually no historians try to increase the probability of making accurate causal statements by using



such well-known techniques as multiple working hypotheses or strong inference.

Recent authors on the history of education have generally relied upon two procedures to arrive at convictions about cause. Both purport to be based upon assessment rather than imputation, and both are commonly used by historians. The first procedure is simply an assertion of a relationship between one sequence of events or one situation and something following it which the historian considers to be obvious. For instance, Beach (1968) portrayed the growing professionalism of college administrators at the turn of the present century. He suggested several reasons for this professionalism and contended that it had identifiable consequences. His argument was based explicitly on the premise that those consequences were a result of the professionalism and not simply simultaneous with it. Further, he maintained this series of events is still in process, and therefore must be taken into account in any contemporary analysis. Cruse's (1967) depiction of twentieth-century Negro intellectual life was written with the same purpose in mind. Both Beach and Cruse based their contentions on reasoning essentially similar to that which underlies the notion of face validity.

Feuer (1969) is an example of the second type of statement about cause, a type that rests upon an assumption of the similarity of the human condition between cultures and across time. He argued that a generation gap occurs in all societies subject to certain social conditions, that this gap at its most extreme leads to the "de-authorization of the elder generation," and that one result is what is known as a "student movement." Feuer's work is based upon three premises: first, that humans everywhere react pretty much the same way to the same conditions; second, that student movements have all the characteristics of mass Oedipal complexes; and third, that societies can avoid some of the "irrationalities and self-destructive components" of student movements by becoming more consciously aware of them. Satisfied with these premises, Feuer brilliantly traced the history of student movements from Leningrad to Berkeley, Athens to Tokyo.

### *Drawing Conclusions and Making Implications*

Most historians assume without question that humans in various times and places are essentially the same. They attest to that assumption each time they rely upon empathy as the sole ground for stating conclusions. The historian tries to think and feel like the persons he is studying, yet knows he may have both more facts and greater insight than they did. Virtually any work of history will include conclusions which the author wishes to be accepted on the basis of sheer faith in his knowledge. Such conclusions are necessary to give the work its sense of drama, as in Rosenberg (1968); its ring of authenticity, as in Tyack (1967b); or its entitle-



ment to human interest, as in Sparrow (1967). The best recent example of the unquestionable appropriateness of face or imputed validity is Wilkinson's (1964) elegant portrait of the values developed in the British public school and their relationship to British political life.

Face validity and empathy have an important characteristic in common. Frequently they are the only means at the historian's command for making judgments about cause. When this is true, the researcher must be particularly careful to avoid "presentism," or the imposition of modern patterns of thought upon the minds of a different era. Avoiding presentism is often equated with being objective in historical research. This has proven particularly difficult for historians concerned with education. There are two reasons. First, they tend to be keenly aware of current educational practices and policies and to have strong views about them. They are tempted to make their history serve these views. This was the basic criticism leveled at the first generation of textbook writers in the field by Bailyn (1960) and Cremin (1965). But Katz (1968a and b) is just as guilty as that previous generation of tethering history to a vision of the schools.

The second reason historians of education are often seduced into presentism is their flirtation with sociological and psychological concepts and tools of analysis. This growing love affair is not unique to historians of education, but is very apparent in much of their recent writing. Consider the explicit concern of Katz (1968a) and Tyack (1967a) with bureaucracy, of Karier (1967) and Wilson (1968) with community, of Bowers (1967) and Fisher (1967) with ideology, of Ambrose (1966) and Wiebe (1969) with institutions, of Beach (1968) and Calvert (1967) with professionalism, and of Feuer (1969) and Joncich (1968) with Freudian views of father-son relationships. The mere decision to use one or more of these constructs represents a value judgment that too often goes unexamined. Concepts from the social sciences help historians think about historical phenomena and make them more intelligible to their readers. But the social sciences themselves are of recent vintage and may not have yielded appropriate guides for thinking about the past. For example, concepts such as bureaucracy and stratification may not be germane to an essentially agrarian society. For an extensive commentary on historians' use of concepts from psychology, see Little (1969).

The intentional use of these organizing ideas makes clear the chief reason for the counter-factual nature of much recent historical writing. Counter-factual arguments are contentions that events or situations would have turned out differently in the absence of some specific factors that did exist. For most historians counter-factual reasoning is an inadvertent corollary to their assertions about cause. It is, therefore, usually implicit rather than explicit. Although such analyses could be consciously made by

imaginative applications of game theory to historical research, they are most commonly found in accounts of phenomena whose consequences the historian thinks were unfortunate. For example, Katz (1968b) maintained that the goals of school reformers in periods of urbanization have precluded in this country a vision of the public school as an institution bringing "joy and delight to the life of the individual." His analysis may well be correct. The point is that it carries with it an "if-then" supposition. Katz implied that, in the absence of such goals, the vision and therefore the nature of our public schools would have been different. Counterfactual implications are inherent to statements about cause. Thus, they will be found to some extent in virtually all of the works discussed in this essay. For a more detailed discussion of counter-factual statements, see Fogel (1967).

### *Conclusion*

The foregoing remarks reveal three main trends in recent research on the history of education. The first is that inquiry remains focused essentially on schooling to the exclusion of other social influences and agencies purposefully designed to change behavior. This inquiry, is however, becoming more sophisticated in response to the second trend, an increasing use of concepts from the social sciences. The most frequently applied concepts center on the schools as social institutions and include analyses of schools as bureaucracies and purveyors of ideologies.

The third trend is the persistent concern with education in the nineteenth century. This should not be surprising, for the field was first explored by men who were convinced that the major battles for public education had been won during that period. That conviction influences current research more than historians realize. Historical research in the nineteenth century is tempting also from the standpoint of available data. Men of that era still had the need and the leisure to put their thoughts on paper. Materials are rich and well-catalogued. In contrast, the twentieth century has not proven so attractive to the historian. The fault lies partially with the changing nature of data. Easier travel and the telephone tend to make conversation the medium of private thoughts. In addition, our society has shown an increasing passion for collecting numerical data and has, therefore, confronted its historians with historiographic problems unanticipated in their graduate training. Of course, technology itself could become a servant rather than a threat to historians. Indeed, some are learning to use computers and other such aids. A few others are showing some interest in oral history, the making and use of tape recordings of interviews with important contemporary figures. But virtually no historians have even begun to develop techniques for using photographs and other visual material as data. The same can be said for artifacts such as materials

for science demonstrations, products of shop and sewing rooms, and school buildings themselves. In general, the imaginations of most historians remain limited to the printed or cursive page. Few have displayed the gift for seeing evidence in pictures and stories so delightfully revealed in the classic history of childhood by Aries (1962).

The reluctance to enter seriously into the study of education in twentieth century America has resulted in neglect of several significant questions and continuities. For example, there is nothing even approaching an adequate study of the history of community colleges. Therefore, these important institutions are given only cursory treatment in even the most complete texts. Yet they represent much that is both good and evil in American education: its concern for expanded opportunity and community involvement, and its petty and impersonal treatment of students. Furthermore, the relationship between community colleges and other social agencies—police and fire departments, the professional aspects of medicine and dentistry, and the management of leisure in such places as motels and restaurants, for instance—makes the community college an ideal focus for the study of social history at the local level. Finally, the development and significance of adult education is reflected in much of the work of these fledgling institutions.

The education of adults on a large scale has yet to intrigue historians. This is especially unfortunate because of the potential lessons to be learned from this process. There are few enterprises so varied. Consider the training programs of business and industrial firms, military organizations, the Foreign Service and other agencies of government at all levels. Consider also the variety of purpose in adult education. Students engage in their studies to learn survival techniques, to raise their salaries by banking credits, to refresh themselves intellectually and artistically, to find or avoid a spouse, and to replace an outdated skill with one of greater market-value. From still another perspective, adults are subjected to the widest conceivable variety of teaching and motivation techniques. For all these reasons, the history of adult education is the history of much that is innovative, imaginative, and humane in the learning process.

It is not surprising that historians have spurned the study of adult education, for it does not fit conveniently into the categories they have established for their work. Adults may learn in school buildings, but they stopped being schools when the children left for the day. Adults may learn informally, with no curriculum theorist or professional society to hail their success. In other words, to study adult education is to forego an assumption implicit in much educational research: education is a verb, not a noun. It is action with people designed to change their behavior in a given direction. Because of that assumption historians study the initiators of that action, their behavior and their products, and the institutions



surrounding and supporting them. As Bailyn (1960) pointed out, historians have limited themselves within the confines of this assumption to the study of schooling. As a result there are very few studies of even the scope of Machlup's (1962), in which he considered a wide range of agencies by which knowledge is produced and diffused. The historian's typical approach is to examine contrasting statements about the ends and means of education and to assume those that appear most official were also most influential. Few historians have investigated the results of schooling. Fewer still have tried to trace these results to their causes. Most writing about the history of education, like reports of most other research in education, rests upon the assumption that intentions are equivalent to results. This propensity is no better illustrated than in Bidwell's (1967) paper. Bidwell placed various kinds of data into categories which could be statistically manipulated. He then drew conclusions about moral education based upon those manipulations. The value of the paper lies in new ways of handling standard data, not in using new evidence. The research itself was predicated on the assumption that pupils learned what their elders intended.

If historians could escape the assumption that education is something done to people rather than, for instance, something they seek and gain on their own, there might be more studies with the extraordinary perception and sensitivity of Cruse (1967). Like Machlup, Cruse was concerned with many agencies of behavior change. He discussed them, however, from the standpoint of the receiver rather than the giver. Cruse presented the artistic and intellectual milieu of Harlem as educative in the sense that it was a condition which shaped the behavior of people who partook of it. In the same vein, there are larger aspects of American culture which profoundly influenced schooling and education, but which have gone relatively unexamined. Some have been pushed to the sidelines by research on themes that appear more important. The concept of industrialization, for example, is a useful shorthand for an array of relatively new factors on the social and political landscape. But in their zeal to plumb its significance, historians have ignored the forces of rural life that continue to pervade American culture. Mentioning these forces should not call to mind only gerrymandered states and forsaken ghettos, but also the emigration of rural values to the city. Thus educators continue to be concerned with general culture rather than special skills and to show growing interest in educational parks and campuses.

There are two other perspectives on the topic of milieu which should be more commonly reflected in the work of historians of education. One concerns the cultural setting, the other the professional climate of opinion. With respect to the first, historians have hardly begun to take full advantage of comparative study as a technique for analysis. This review has included comments on two books whose authors make bows in this direc-



tion. Jencks and Riesman (1968) described different kinds of institutions in this country. Feuer (1969) portrayed student movements in various nations. Like much of the writing in comparative education and comparative history, however, their books require the reader to do most of the comparing. Consider how Fisher (1967) might have handled her study of industrial education. There are few cultures whose systems of education, whether formal or informal, have not included some notion of the relationship between productive labor and concepts of human nature, good citizenship, and happy living. A comparison between various American visions of industrial education and the changing definitions of polytechnic education in the U.S.S.R., for example, might have proven more instructive than limiting her inquiry to one industrial society. As another example, Calhoun (1966) might have compared the role of the nineteenth century American intellectual with that of intellectuals during the formative years of other English speaking societies such as Canada and Australia. This is not to say that historians need run the risk of reasoning by analogy, but only that they may develop larger insights by a somewhat more comprehensive view of what they hope to understand.

Many of the foregoing remarks have been comments upon the professional climate of opinion within which research on the history of education is carried out. One further comment is in order. Historians of education, like most academics, spend virtually no time studying either their own discipline or education itself as a professional endeavor. The result is a disturbing paucity of studies on some of the most perplexing and important questions faced by educators in America. These questions include such concerns as the internal politics of institutions, the growth of academic disciplines, and the relationship between educational research and schooling. To ignore such matters is to ignore inquiry with data intimate to our daily lives and with outcomes potentially critical to society.

Historians of education are in particularly advantageous positions to examine questions such as these. Usually their academic homes are schools of education rather than departments of history. Thus, their colleagues and their students, as well as the data for their research, represent a wider variety of intellectual approaches to a common set of problems than is found in most other fields of inquiry. Further, they are, to use Bruner's (1962) suggestive concept, part of the left hand of the enterprise of educational research. Their work is still more art than science and, like the practice of education itself, likely to remain so. They must do better at interpreting the enterprise they share with their colleagues.

### *Bibliography*

- Ambrose, Stephen E. *Duty, Honor, Country: A History of West Point*. Baltimore, Md.: The Johns Hopkins Univ. Press, 1966.  
 Aries, Phillippe. *Centuries of Childhood*. New York: Knopf, 1962.

- Bailyn, Bernard. *Education and the Forming of American Society*. New York: Random House, 1960.
- Barnard, John. *From Evangelism to Progressivism in Oberlin College, 1866-1917*. Columbus: The Ohio State Univ. Press, 1968.
- Beach, Mark. Professional versus Professorial Control of Higher Education. *Educational Record* 49: 263-73; 1968.
- Benson, Lee. An Approach to the Scientific Study of Past Public Opinion. *Public Opinion Quarterly* 31: 522-67; 1968.
- Bidwell, Charles. The Moral Significance of the Common School. *History of Education Quarterly* 6: 50-91; 1966.
- Bowers, C. A. The Ideologies of Progressive Education. *History of Education Quarterly* 7: 452-73; 1967.
- Bowers, C. A. *The Progressive Educator and the Depression: The Radical Years*. New York: Random House, 1969.
- Bragdon, Henry W. *Woodrow Wilson: The Academic Years*. Cambridge, Mass.: Belknap Press of Harvard Univ. Press, 1967.
- Bruner, Jerome. *On Knowing: Essays for the Left Hand*. Cambridge, Mass.: Harvard Univ. Press, 1962.
- Bullock, Henry A. *A History of Negro Education in the South from 1619 to the Present*. Cambridge, Mass.: Harvard Univ. Press, 1967.
- Calhoun, Daniel. From Noah Webster to Chauncey Wright: The Intellectual as Prognostic. *Harvard Educational Review* 36: 427-44; 1966.
- Callcott, George H. *A History of the University of Maryland*. Baltimore, Md.: Maryland Historical Society, 1966.
- Calvert, Monte. *The Mechanical Engineer in America. Professional Cultures in Conflict*. Baltimore, Md.: The Johns Hopkins Univ. Press, 1967.
- Clubb, Jerome and Howard, Allen. Computers and Historical Studies. *Journal of American History* 54: 599-607; 1967.
- Cremin, Lawrence. *The Transformation of the School*. New York: Knopf, 1961.
- Cremin, Lawrence. *The Wonderful World of Ellwood Patterson Cubberley: An Essay on the Historiography of American Education*. New York: Teachers College Press, 1965.
- Cruise, Harold. *The Crisis of the Negro Intellectual*. New York: Morrow, 1967.
- Daniels, George. *American Science in the Age of Jackson*. New York: Columbia Univ. Press, 1968.
- Drost, Walter H. *David Snedden and Education for Social Efficiency*. Madison: Univ. of Wis. Press, 1967.
- Feuer, Lewis. *The Conflict of Generations: The Character and Significance of Student Movements*. New York: Basic Books, 1969.
- Fisher, Berenice M. *Industrial Education: American Ideals and Institutions*. Madison: Univ. of Wis. Press, 1967.
- Fishlow, Albert. Levels of Nineteenth Century American Investment in Education. *Journal of Economic History* 26: 418-36; 1966a.
- Fishlow, Albert. The American Common School Revival, Fact or Fancy? *Industrialism in Two Systems*. (Edited by Henry Rosovsky.) New York: John Wiley and Sons, 1966b.
- Fogel, Robert. The Specification Problem in Economic History. *Journal of Economic History* 27: 283-308; 1967.
- Gardner, David P. *The California Oath Controversy*. Berkeley: Univ. of Calif. Press, 1967.
- Graham, Patricia Albjerg. *Progressive Education: From Arcady to Academe. A History of the Progressive Education Association, 1919-1955*. New York: Teachers College Press, 1967.
- Hayter, Earl W. *The Troubled Farmer, 1850-1900: Rural Adjustment to Industrialism*. DeKalb: Northern Ill. Univ. Press, 1968.
- Jencks, Christopher and Riesman, David. *The Academic Revolution*. Garden City, N.Y.: Doubleday, 1968.
- Joncich, Geraldine. *The Sane Positivist: A Biography of Edward L. Thorndike*. Middletown, Conn.: Wesleyan Univ. Press, 1968.

- Karier, Clarence. *Man, Society, and Education: A History of American Educational Ideals*. Glenview, Ill.: Scott, Foresman, 1967.
- Katz, Michael. The Emergence of Bureaucracy in Urban Education: The Boston Case, 1850-1884. *History of Education Quarterly* 8: 155-88, 334-56; 1968a.
- Katz, Michael. *The Irony of Early School Reform: Educational Innovation in Mid-nineteenth Century Massachusetts*. Cambridge, Mass.: Harvard Univ. Press, 1968b.
- Kett, Joseph F. *The Formation of the American Medical Profession. The Role of Institutions, 1780-1860*. New Haven, Conn.: Yale Univ. Press, 1968.
- Kevles, Daniel. Testing the Army's Intelligence: Psychology and the Military in World War I. *Journal of American History* 55: 565-81; 1968.
- Lannie, Vincent. *Public Money and Parochial Education: Bishop Hughes, Governor Seward, and the New York School Controversy*. Cleveland, Ohio: Press of Case Western Reserve Univ., 1968.
- Little, Lester K. Psychology in Recent American Historical Thought. *Journal of the History of the Behavioral Sciences* 5: 152-72; 1969.
- Machlup, Fritz. *The Production and Distribution of Knowledge in the United States*. Princeton, N.J.: Princeton Univ. Press, 1962.
- Parsons, Kermit. *The Cornell Campus: A History of its Planning and Development*. Ithaca, N.Y.: Cornell Univ. Press, 1968.
- Rader, Benjamin. *The Academic Mind and Reform: The Influence of Richard T. Ely in American Life*. Lexington: Univ. of Ky. Press, 1966.
- Rosenberg, Charles. *The Trial of the Assassin Guiteau*. Chicago: Univ. of Chicago Press, 1968.
- Rothblatt, Sheldon. *The Revolution of the Dons: Cambridge and Society in Victorian England*. London: Faber and Faber, 1968.
- Smith, Timothy. Protestant Schooling and American Nationality, 1800-1850. *Journal of American History* 53: 679-95; 1967.
- Sparrow, John. *Mark Pattison and the Idea of a University*. Cambridge: Cambridge Univ. Press, 1967.
- Stambler, Moses. The Effect of Compulsory Education and Child Labor Laws on High School Attendance in New York City, 1898-1917. *History of Education Quarterly* 8: 189-214; 1968.
- Tyack, David. Bureaucracy and the Common School: The Example of Portland, Oregon, 1851-1913. *American Quarterly* 19: 475-98; 1967a.
- Tyack, David. *George Ticknor and the Boston Brahmins*. Cambridge, Mass.: Harvard Univ. Press, 1967b.
- Tyack, David. The Perils of Pluralism: The Background of the Pierce Case. *American Historical Review* 74: 74-98; 1968.
- West, E. G. The Political Economy of American Public School Legislation. *The Journal of Law and Economics* 10: 101-28; 1967.
- Wiebe, Robert H. The Social Function of Public Education. *American Quarterly* 21: 147-64; 1969.
- Wilkinson, Rupert. *Gentlemanly Power: British Leadership and the Public School Tradition*. London: Oxford Univ. Press, 1964.
- Wilson, Raymond J. *In Quest of Community: Social Philosophy in the United States, 1869-1920*. New York: Wiley, 1968.
- Wishy, Bernard. *The Child and the Republic: The Dawn of Modern American Child Nurture*. Philadelphia: Univ. of Pa. Press, 1968.

## AUTHOR

MARK BEACH Address: University of Rochester Title: Associate Dean, College of Arts and Sciences Age: 32 Degrees: B.S., Univ. of Wis.; M.A., Univ. of Minn.; Ph.D., Univ of Wis. Specialization: History of Education.

### 3: AESTHETIC INQUIRY

DAVID W. ECKER

*New York University*

THOMAS J. JOHNSON

*Central Midwestern Regional  
Educational Laboratory, Inc.*

EUGENE F. Kaelin

*Florida State University*

Whether experimental and psychological or analytic and philosophical, aesthetic inquiry is concerned with isolating a field of relevant data, interpreting it, and when most ambitious, with formulating hypotheses to explain the law-like reoccurrences that may appear in the phenomena under investigation. It is our purpose in this chapter to examine the methodological assumptions and the practical results of applying specific research techniques to "aesthetic" phenomena in order to determine the workability of educational procedures that depend upon various methods of aesthetic research. We examine the claims of psychologists, linguistic philosophers, and phenomenologists that they have the methods appropriate to develop a workable tool for teaching aesthetic materials.

#### *Experimental psychology*

Two dominant experimental techniques for determining the aesthetic preferences of individual subjects—rank ordering of items and paired comparisons—stem from the simple "method of choice" introduced into psychology by Gustav T. Fechner (1876). Fechner instructed his subjects to choose which one of a given number of rectangles was more pleasing to that person's taste. (See Woodworth and Schlosbert, 1954, pp. 252-61.) When, instead of giving a simple choice, subjects are asked to rank objects from most to least preferred, there is an obvious increase in the amount of information obtained in "choice" experiments. Still more information should be obtained by requiring the subjects to choose one of a pair selected from the range, and then exhausting the number of possible pairs to be combined from the items within that range. This last technique has the added advantage of allowing a test of consistency in the individual's



preference judgments. For example, if a subject states his preference for A over B and of B over C, then presuming preference to be logically transitive, the same subject ought to declare a preference for A over C. If the subject states his preference for C over A, then this reversal of the expected judgment produces what has come to be called a "circular triad" (Pratt, 1956).

The frequency of such circular triads then becomes a measure of the inconsistency in the subject's aesthetic preference. Pratt (1956) reported a study of 178 high school students who were asked to state their preferences among colors, designs, and written opinions in six different series of stimuli. In each series 12 stimuli were presented by the technique of paired comparisons, thus requiring 66 judgments from each subject or 11,748 judgments from 178 subjects. The mean number of circular triads actually made by the 178 subjects in each series was only 6.25 out of seventy triads possible, with a minimum of 37 attributable to chance. Whether or not the claim may be made to stick that preference is a transitive relation, the facts indicated a high degree of consistency of preference and led Pratt to maintain that "logic" does indeed apply to a set of phenomena popularly believed to be immune to the probes of objective analysis.

Perhaps the most comprehensive studies of aesthetic preferences in recent times were those of Irvin L. Child (1962, 1964, 1967). In his *Development of Sensitivity to Aesthetic Values* (1964, p. 1), Child stated that "aesthetic sensitivity may be measured if people's responses to standardized presentation of works of art can be evaluated against a suitable criterion." The criterion selected was the convergence in the preferences of experts, who presumably are capable of stating reasons in support of their own aesthetic judgments.

Although it is obvious that Child begged the issue of the validity of experts' judgments (for *their* judgments to be tested one would need another panel of experts; and for theirs, another; and so on, *ad infinitum*), he did select a number of parameters that he considered to determine expert judgments. These parameters are selectivity; organization; consistency in use of line, color, shape, texture, etc.; "aliveness" or vitality, suitability of decoration; and "appropriateness" in the use of elements. After training school children to recognize these elemental aesthetic properties of works of art, Child found that their judgments were improved. That is, given a clear understanding of the reasons why experts tended to prefer one work over another, these subjects could recognize or guess which of two items the experts preferred.

In a second experiment, however, Child (1964) found that mere continued exposure to many works was insufficient to enable students' preferences to have any higher degree of consistency with those of experts. In

other words, continuous exposure to the pairing without explicit statement of aesthetic principles of judgment was insufficient to overcome the effects of boredom, fatigue, personal prejudice, etc. It should be remembered, nevertheless, that in the second instance the subjects were not told to try to give the "correct" answer, although their preferences were ultimately compared to the judgments of experts.

Valentine (1962) conducted a survey of the characteristic research of the last fifty years. He examined studies of attitudes and preferences in the perception of colors, lines, and shapes, as well as experiments with complete organized paintings; and his interest was broadly enough defined to include the results of work done in music and poetry. Also interesting are the studies of Leijonhielm (1967) and Eysenck's chapter on the psychology of aesthetics, in his revealing *Sense and Nonsense in Psychology* (1957).

The task here is to separate sense from nonsense in applying psychometrics to aesthetic responses. Child (1964) distinguished between preference and judgment. Clearly, for the panel of experts there is and can be no such distinction. Child's operational definition of judgment for inexpert subjects simply cannot be applied to any one of the theoretically infinite panels of expert judges whose preferences are unquestionably accepted as standard aesthetic judgments. Experts are experts, presumably because they know and can state reasons for what they like. The convergence in their preferences is the criterion of "aesthetic value."

Consider the following linguistic evidence: one expert declares, " $x$  is a better painting than  $y$  for reasons  $a$ ,  $b$ , and  $c$ ; but I prefer  $y$ ." Since there is no obvious contradiction in this declaration, it leads the reader to believe that the relative goodness of the works in question has nothing to do with *unqualified* preference. Since one can like a given painting for any reason, the problem is to determine the aesthetic relevance of the reasons offered in support of preferences (Ecker, 1967). These reasons should not be equated with "causes." The host of correlational studies linking personality traits with creativity or aptitude for aesthetic judgment, even if predictive of future student performance, obviously does not explain the relevance of a subject's reasons for preferring works of art or the manner in which he actually responds to aesthetic stimuli.

The value of experimental studies of aesthetic preferences is, therefore, limited to the degree that, for an expert at least, a preference can be given a reason: that is what is meant by "aesthetic judgment." The results of Child's study (1964) demonstrated that limitation. In the only positive results he recorded, his subjects were previously trained to recognize the characteristics cited by critics as reasons for preferring works of art. When this prior training was not given, the subjects' preferences did not show any shift toward or away from agreement with expert preference. Thus,

Child's success depended upon the short course in aesthetic discrimination that he gave his subjects, but not upon the technique of paired comparisons for determining their preferences.

All preferences entail comparisons, but one does not compare two works of art to determine individual goodness. Rather, one judges each work of art individually in order to determine relative goodness (Ecker, 1967). This distinction between individual and comparative judgments becomes masked in the notion of a "preferential judgment"; and this same distinction may turn out, in further inquiry, to constitute one of the necessary conditions for measuring the changes in aesthetic responses, either by the researcher, in his subjects; or by teachers, in their students; or by the student, in his own progress towards fuller aesthetic knowledgeability. The more knowledgeable one becomes, the more one is capable of judging a single work of art.

*Additional References:* Barron (1952); Beebe-Center and Pratt (1930); Beittel (1964); Birkhoff (1932); Child (1962, 1966); Davis (1936); DeWare (1933); Eisner (1966a, b); Eysenck (1941a, b, c, d); Getzels (1964); Hambridge (1926); McWhinnie (1968c); Wilson (1966).

### *Analytical philosophy*

Professor Child's success with developing sensitivity to aesthetic value was clearly dependent upon a linguistic phenomenon; he succeeded in formulating a set of aesthetic categories that were empirically adequate to refer to the properties of good works of art. His choice of language for naming and explaining these categories was, of course, guided by the level of understanding of elementary school children; the categories themselves apparently were derived from "... some of the ideas art critics express in talking about the relative aesthetic value of different works" (Child, 1964, p. 36).

In light of the foregoing distinction between individual and comparative judgments, however, a second look at talk about art and at the critical process by which individual aesthetic judgments are made and justified (whether by art critics, teachers, or children) seems appropriate. The determination of categories for understanding art and, more recently, the clarification of the language of art criticism constitute the major work of analytic aesthetics; it is precisely at this level of "meta-criticism" that a fruitful relation may be constructed between researcher, teacher, and student.

Monroe C. Beardsley is one of the most influential of contemporary analytic philosophers who define the domain of their concern for aesthetics as the language used by critics in evaluating works of art. Beardsley's principal work (1958) is sub-titled "Problems in the Philosophy of Criticism."



After examining the great variety of statements made about works of art, Beardsley devised a set of seven postulates to differentiate between adequate and inadequate statements about the nature of an individual object. These postulates are one philosopher's attempt to stipulate the conditions for evaluating precisely what Child, as an experimental psychologist, felt obliged to leave undone; namely, to assess the appropriateness of critical language to encapsulate the experience afforded by the perception of an individual work of art.

For example, in postulate one, Beardsley differentiated between the aesthetic object *qua* perceptual and its many presentations (or appearances), which, according to the second postulate, may occur at different times with different people. In postulate three, he admitted the possibility of two presentations of the same object being different from each other, and in postulate four, he acknowledged the possibility that a single presentation may reveal less than the totality of characteristics displayed by any given aesthetic object. In postulates five, six, and seven, Beardsley attempted to stipulate the conditions for determining the difference between "veridical" and "illusory" presentations. A presentation is veridical when its characteristics correspond to those of the aesthetic object; it is illusory when they fail to do so. Lastly, in postulate seven, he stated that when two presentations have incompatible characteristics at least one must be illusory.

For researchers who would use Beardsley's postulates, the problem, which in experimental aesthetics may be considered insoluble, occurs in determining the difference between "the aesthetic object" and its manifold "presentations." Philosophers criticize each other for appealing to a distinction between "the physical object," such as pigments on canvas or sound waves, and the "aesthetic object" which appears only to the attentive consciousness. Ziff (1954) maintained that such a distinction is spurious because the only difference between the so-called physical object and the "object of art" is a difference in linguistic description and not in kinds of (metaphysical) objects. His point is well taken; people rarely become confused when speaking about an art object in strictly physical terms. A painting may cover a 4' x 6' canvas, but the space tensions created by its colored masses are not capable of physical measurement. In experimental terms this difference is one of considering the stimulus (e.g., color) in an absolute sense from another which is relative to the psychological processes of the perceiving organism, in whose responses the physical stimulus becomes suffused with the warm character of human affectivity (qualitative response).

Value-toned perception was the definitive characteristic isolated by the first American aesthetician, George Santayana (1896), when he suggested that the domain of aesthetic inquiry be delimited by broadening criticism to include immediate apprehension of aesthetic characteristics (not just an



explicit referral to a critical canon) and by narrowing the notion of perception to include only those stimuli which "objectify" human feeling. It is not sufficient, therefore, merely to differentiate the kinds of language used to determine relevant critical sense. It is necessary at a higher methodological level of inquiry to stipulate rules for correctly using a language of criticism to disclose the nature of an aesthetic object—one that has engaged human appreciative faculties.

Beardsley's postulates constitute at least the first step in the attempt to differentiate critical sense from nonsense. His postulates, of course, were intended for the descriptive use of the term *art* and not for its normative use. However, if the normative judgment of works of art is to be defined in terms of human liking or preference, it seems fair to start with an adequate description of what it is exactly that an individual subject prefers. Therefore, a student should not be asked whether he likes a given object or prefers it to another, he should be asked the following question: "What is there in the experience of this object to be liked?" Giving an answer to this question will allow anyone, including students, to state reasons for his preferences. This is what Child's "short course in aesthetics" accomplished, however imprecisely, in allowing the subjects to enhance their sensitivity towards aesthetic value.

In general a work of art, descriptively considered, is a candidate for human appreciation; a good work of art is one for which the appreciator can stipulate his reasons, from an analysis of the work, for claiming that his attention has been fruitfully engaged. Thus, Weitz (1956) pointed out that there is no definite set of necessary and sufficient conditions for the goodness of aesthetic objects, but only congeries of changing, adventitious characteristics as new excellencies are created and discovered. He thereby frees the philosopher from the age-old and unachieveable objective of determining "essences."

More systematic attempts of linguistic aesthetics are found in the work of Morris (1939a, 1939b, 1957) and Langer (1948, 1953, 1957). Aesthetics considered either as a general theory of linguistics or as a special applied branch of general linguistics is found in the work of Croce (1909), Collingwood (1958), Merleau-Ponty (1952), Ziff (1960), and Goodman (1968). Margolis (1965) presented a convenient summary of the most outstanding issues analytical philosophers have been developing in aesthetic theory; his work is aptly entitled *The Language of Art and Art Criticism*.

In scanning the attempts of linguistic philosophers to unravel some of the muddles attendant upon critics' loose use of critical terminology, non-analytical philosophers have pounced upon the remoteness of linguistic-analytic inquiries from the description of the concrete conditions under which individual works of art are experienced and enjoyed. Croce (1909) defined the "aesthetic" as knowledge of the radically individual objects of

sensuous intuition, clearly to be distinguished from concepts, which are the objects of scientific knowledge. If Croce is correct, this distinction between intuitive, or aesthetic, and conceptual, or scientific, knowledge would explain why any concepts used in the "science" of aesthetics must always be of a second or methodological order which, to be empirically adequate, must be rooted in valid descriptions of first-order sensuous intuitions. (See Munro, 1956.)

The total analytic process would therefore entail the following phases: (a) experience of the individuality in an aesthetic expression, (b) description of the structures inherent within the experienced objects, and (c) isolation of repeatable features which may be shared by many works of art. Only such features are properly called "aesthetic categories."

Failure to heed the distinction of orders or levels at which language is operative enabled Child (1964) to equate preference with judgment; on the other hand, over-zealousness on the part of linguistic philosophers to heed the same distinction made it impossible for them to indicate the connection between experience of art as (say) the language of vision (first-order intuitive response to organized visual surfaces) and the aesthetic categories or concepts used to describe the way in which the "visual" language comes to operate within aesthetic communication. The problem, of course, is for the aesthetic inquirer to devise a set of categories capable at least of calling attention to the phenomenal characteristics of aesthetic objects.

*Additional References:* Beardsley and Schueller (1967); Ecker (1966); Gallie (1948); Margolis (1965); Pepper (1956); Prall (1967; Smith (1966); Ziff (1958).

### *Phenomenological analysis*

Ordinary language is quite capable of assigning names to aesthetic categories. The "plain man" seems satisfied with the notions of matter, form, and feeling. He needs only a critical examination of these particular categories to determine how each may be referred to individual works of art. The only stipulation we should like to add is that some procedure must be devised for distinguishing relevant from irrelevant statements about the nature of the aesthetic object in question whenever the plain man applies these categories to his experiences.

American phenomenologists, in particular Kaelin (1970), have interpreted Husserl's *epoché* (1931) as a technique for establishing relevance in statements of criticism. As Husserl developed the technique, practicing the "phenomenological *epoché*" entailed the suspension of what he called the "natural attitude." This latter term is used to refer to the life conditions of ordinary humans relating to the objects of their natural environment, interpreted either through the categories of common sense or of scientific ex-

planation. Thus, when a subject looks around and says that he sees a tree or describes the conditions conducive to the health of that tree, he testifies to the natural attitude within which most people live. To suspend this attitude, one merely attends to the qualities of the object in question as they appear to one's conscious attention. Therefore, by "bracketing out" all the non-phenomenal characteristics of the appearances of the natural object, such as its species, any of its supposed causes, or putative explanations associated with its existence, one is left with what John Dewey (1934) referred to as the "felt quality of the immediate."

We are not maintaining here that there is a mysterious aesthetic attitude according to which all aesthetic properties of a given physical object automatically become revealed to a viewer. Dickie (1966) criticized this naive notion of an aesthetic attitude as he found it expressed in the aesthetic theory of Aldrich (1965), whose *Philosophy of Art* is portrayed by its author as a "phenomenological" approach to the determination of aesthetic values. Dickie's charge has some merit: Aldrich admitted that his work is not in the strict Husserlian tradition, and there is no royal road to the experience of aesthetic quality.

The function of the *epoché* is to close off all irrelevancies that may occur to an appreciator as he attends to the qualitative base of the aesthetic expression before him. To practice the *epoché*, one is enjoined from referring to the properties of the object *qua* physical or to any other associations which his perception may invoke, but which are not controlled by the qualitative structure he perceives.

The advantage of the *epoché* is to avoid distinguishing, as Beardsley has done, between "veridical" and "illusory" presentations. In the strictly phenomenological scheme, the bracketed appearances are the aesthetic object. As Ziff (1954) maintained, there are not two kinds of metaphysical objects; there is, however, a single experience in which the phenomenal characteristics of a single object are set off, as if "in brackets." Within these brackets occurs the totality of a field of relevance, referred to as "the context of aesthetic significance." The analytical task of the theoretician of art is to devise a set of categories capable of describing how significance accrues within this context.

Kaelin (1964, 1968a, 1968b) attempted this task and, like Beardsley, formulated his conclusions in a set of interpretive postulates. They are: P1. All aesthetic significance is context bound. P2. An aesthetic context is composed of "counters" and their relations, either surface only or surface and depth. P3. No aesthetic counter has absolute significance; i.e., each has only that significance which is made apparent by a relationship to some other counter within context. P4. The significance of the context is the felt expressiveness of all the counters as they "fund" or come to closure in a single experience. By "counters" is meant anything that is discriminable in the

bracketed context. Surface counters are sensuous stimuli; and depth counters, anything of a cognitional nature (i.e. conceptual or representational elements) bound to these surface counters. Works of non-objective art and absolute music, for example, have their significance exhausted in the perception of the relations between surface counters, but works of representational art contain additional depth counters.

Note that by P3 the significance of the work in question is not to be attributed to any single counter or to any isolated set of counters, whether surface or depth. A critical examination of many of the experimental studies in aesthetics, such as those by Fechner (1876) and Child (1964), reveals a violation of this crucial postulate. P3 is also violated in most attempts by inept societies to censor aesthetic expressions for the appearance of an immoral or obscene subject matter. As long as investigators continue to measure feelings, attitudes, beliefs, value sets, personality traits, and the like, independently of a close correlation with the conditions described in these postulates, whatever else is claimed, there can be no claim to relevance for *aesthetic* education. It may be otiose to note that curriculum planners in the field of aesthetic education must heed this enjoinder to aesthetic relevance, but when properly performed, phenomenological aesthetics effectively becomes a philosophy or methodology of aesthetic education.

The function of language in the philosophy of aesthetic education is first of all categorical; that is, it must devise a workable set of aesthetic or critical notions such as those found in any of the three sets of categories cited above (Child, 1964; Beardsley, 1958; Kaelin, 1964). Language also functions in the various subjects' attempts to describe what happens as they experience works of art. These descriptions, which are of the individual expressiveness controlled by a given work of art in our first-order experiences, may be called "phenomenological." They are both fallible and corrigible, with reference to similar statements made by others; they may be expressed in any ordinary language, whether that language is used by a theoretician, an experimenter, a teacher, a subject, or a student.

Only one caution is necessary: whatever truth is contained in these descriptions must be measured against an authentic experience of the relevant counters. This proviso corresponds with the technique exploited by Wertz (1956), who explained the philosopher's task as the elucidation of the meaning of aesthetic concepts; by showing how such concepts are used to refer to the phenomenal characteristics of aesthetic objects one would have gauged whatever validity they possess.

*Additional References:* Kaelin, 1962, 1968; Stumbo, 1968.

### *Aesthetic education*

Under the direction of T. J. Johnson, the staff of the Central Mid-western Regional Educational Laboratory (CEMREL) is currently compiling



the results of a survey of the empirical aesthetic inquiry published in the past seventy years. The primary objectives of the survey are to present an overview of the "state of the art"; to provide a reliable and useful summary of research findings, knowledge, and tests; and to develop an index to these research studies so scholars, researchers, and others will have ready access to any information relevant to the field. The actual search of the literature was accomplished by systematically reviewing the various educational and psychological indexes and journals from 1900 to the present, as well as from summary reviews and bibliographies of others. To date, approximately 5000 studies in all areas of aesthetics have been examined.

From CEMREL's attempt to review and briefly summarize these research studies, it was observed that the studies contain three main categories of content:

1. That reflecting the important stimulus classification, parameters, or variables within the particular field of specialization, e.g., Baroque music, Shakespearean poetry, etc.;
2. That reflecting the characteristics of individuals which effect how they react to these variables, e.g., intelligence, age, etc.;
3. That reflecting the implicit or explicit response constructs observed in the study, e.g., children's poetry preferences.

A taxonomy or classification system can now be developed for each of the content areas themselves, which would permit a more detailed analysis of the studies to be made.

Most of the research studies are of the following three types:

*Type 1* studies are those in which all three categories of content are represented. In other words, one or more stimulus dimension is varied; subjects are either randomly assigned to treatments, or strata of subjects are utilized; and the variations in one or more of their responses are observed. In *Type 2* studies, the correlation between two or more response constructs is investigated, and the stimulus dimension is not systematically varied. In *Type 3* studies, a description of the relationship between certain subject characteristics and responses is the objective of the study, and no systematic variation of any stimulus dimension is undertaken. There are a few methodological studies which do not readily fit into any of the above types.

When all the studies are thus typed, it becomes possible to construct a detailed taxonomy of the stimulus, response, and subject characteristics which were investigated in each of the areas of aesthetic inquiry noted earlier. The frequency of studies in each cell of the resultant matrix of study types can be noted, and the "sociology" of empirical aesthetic inquiry in each area will become readily apparent; i.e. the gaps and con-

centrations of the studies, as well as the characteristic modes of inquiry in each area will become manifest.

Table I depicts the major stimulus dimensions varied and the response constructs measured for the various areas of aesthetic inquiry. It reveals a number of interesting aspects of aesthetic inquiry during the period surveyed. Within each area one or two stimulus and response dimensions continue to be the dominant focus of the inquiry.

*Table 1*  
Summary of Major Foci of Type 1 Studies  
According to Area of Investigation

<i>Area of Investigation</i>	<i>Total Number of Studies Located (All Types)</i>	<i>Major Stimulus Dimensions Varied (Type 1)</i>	<i>Major Response Constructs Measured (Type 1)</i>
Graphic Arts	1713	Color, Form	Preference, Identification
Dance	149	Type of Movement	Kinesthetic
Drama	309	Technical Aspects of Production	Audience Preferences
Films	203	Theme	Preference, Emotional Responses
Literature	498	Order and Organization Literary Form	Background Knowledge Preference
Music	1735	Pitch	Discrimination, Identification
Creativity	361	Mode of Presentation	Flexibility

It is apparent from examining Table I that most of the researchers manipulated those stimulus parameters that were simple and easy to manipulate, and measured responses that were relatively simple and easy to measure, and that systematic programming was lacking in many areas. It also seems clear that knowledge derived from empirical aesthetic inquiry will have a cumulative impact only when the research community begins to employ strong *conceptual* models and the established procedures of "strong inference."

The term "strong inference" is borrowed from Platt (1969), who noted that the more rapidly moving fields of research are those in which a particular method is systematically taught and applied to the field. The essential features of strong inference consist in the application of the following steps, formally, explicitly, and regularly, to every problem:

1. Devising alternative hypotheses.
2. Devising crucial experiments, with possible alternate outcomes which may exclude one or more hypotheses.
3. Carrying out the experiment "cleanly."
4. Recycling the procedure, making subhypotheses or sequential hypotheses to refine and reduce the possibilities.

According to Platt, this method of inductive or strong inference yields rapid and powerful progress because it provides the minimum sequence of steps for excluding hypotheses and a regular method for arriving at firm inductive conclusions as rapidly as possible. Although it is clear that "strong inference" is another name for the usual procedures of science, Platt's purpose in restating them is that many researchers seem to have forgotten them. Equipment, calculations, and lectures become ends in themselves; scientists "do busy work" and become method-oriented rather than problem-oriented. The business of science is still to solve human problems.

Using Platt's criteria, it must be concluded that the field of empirical aesthetic inquiry is not currently a rapidly moving field of research. For this reason, the authors have performed the foregoing methodological examination of the techniques and methods used by experimentalists, analytical philosophers and phenomenologists. Future research in aesthetics may profitably be conceived as a meld of the various techniques described.

In order to improve the reliability of future psychological experiments on preference or of correlations between other personality characteristics and aesthetic sensitivity, it seems obvious that the naive criterion of agreement with the convergence of expert opinion must be refined by linguistic analysis of the terms used by the reporting experts. This task is admirably fulfilled by contemporary linguistic-analytical philosophers; and to guarantee the relevance of an analytical philosopher's distinctions to the precise conditions under which an individual work of art is experienced and appreciated, some sort of phenomenological analysis of these conditions must be performed. When it is, the theoretician may turn the job over to practicing scientists and teachers.

It should be clear that the authors of this chapter have focused their attention mainly upon the qualities, of various kinds and degrees of adequacy or authenticity, of aesthetic responses. This choice was made on the grounds that if people are to communicate anything through aesthetic experiences, their responses to them must be similar. Thus, should a

teacher or experimenter wish to examine creativity and its processes rather than active appreciation, then he must show how the creator exercises aesthetic judgment in the various stages of the completion of an art work. For an insight into how this might be done, see the work of Ecker (1963) and its criticism by Beardsley (1965).

Ecker's work was inspired by the methodological inquiries of Villemain (1961) and Champlin (1966), whose source was Dewey (1930, 1934). Dewey's distinction between the "aesthetic" and the "artistic" brings out another aspect of aesthetic experiences left untouched in the present review. One does not, for example, experience immediately felt quality only by responding to artificially arranged sensuous surfaces (see Prall, 1967, pp. 30-75). To handle one's responses to the same sort of quality in the perception of natural objects, one need only carefully place the brackets around whatever object of the natural environment has captured his aesthetic sensitivity.

Further inquiry, both methodological and experimental, remains to be performed at the level of aesthetic education research. In this vein, Kaelin (1968a) attempted to outline a systematic approach to aesthetic education based upon the tenets of existential psychology and phenomenological aesthetics; one of his principles—the projectional—stems from the psychology, and three additional principles—autonomy, relevance, and completeness—are based upon his phenomenological interpretation of aesthetic experiences. This work should be compared to that of such specialists in the field of aesthetic education as Broudy (1964), Arnstine (1964) and Smith (1965, 1966).

*Additional References:* Eisner and Ecker, 1966; McWhinnie, 1965, 1968a, b.

## Bibliography

- Aldrich, Virgil. *Philosophy of Art*. Englewood Cliffs, N.J.: Prentice-Hall, 1965.
- Arnstine, Donald. Shaping the Emotions: the Sources of Standards for Aesthetic Education. *School Review* 72: 242-71; 1964.
- Beardsley, Monroe C. *Aesthetics: Problems in the Philosophy of Criticism*. New York: Harcourt, Brace, 1958.
- Beardsley, Monroe C. On the Creation of Art. *Journal of Aesthetics and Art Criticism* 23: 291-304; 1965.
- Broudy, Harry S. The Structure of Knowledge in the Arts. *Education and the Structure of Knowledge*. (Edited by Stanley Elam.) Chicago: Rand McNally, 1964. Pp. 75-106.
- Champlin, Nathaniel L. Methodological Inquiry and Educational Research. *A Seminar in Art Education for Research and Curriculum Development*. (Edited by Edward L. Mattil.) Cooperative Research Project No. ED-010-000. University Park: Penn. State Univ., 1966. Pp. 291-328.
- Child, Irvin L. *A Study of Aesthetic Judgment*. Cooperative Research Project No. ED-002-819. New Haven, Conn.: Yale Univ., 1962.
- Child, Irvin L. *Development of Sensitivity to Aesthetic Values*. Cooperative Research Project No. ED-003-483. New Haven, Conn.: Yale Univ., 1964.



- Child, Irvin L. *Bases of School Children's Aesthetic Judgment and Aesthetic Preference*. Cooperative Research Project No. ED-012-380. New Haven, Conn.: Yale Univ., 1967.
- Collingwood, R. G. *Principles of Art*. New York: Oxford Univ. Press, 1958.
- Croce, Benedetto. *Aesthetic*. (Translated by Douglas Ainslie.) London: Macmillan, 1909.
- Dewey, John. *Qualitative Thought*. *Symposium* 1: 5-32; 1930.
- Dewey, John. *Art As Experience*. New York: Minton, Balch, 1934.
- Dickie, George. *Attitude and Object: Aldrich on the Aesthetic*. *Journal of Aesthetics and Art Criticism* 25: 89-91; 1966.
- Ecker, David W. *The Artistic Process As Qualitative Problem Solving*. *Journal of Aesthetics and Art Criticism* 21: 283-90; 1963.
- Ecker, David W. *Justifying Aesthetic Judgments*. *Art Education* 20: 5-8; May 1967.
- Eysenck, H. J. *Sense and Nonsense in Psychology*. Baltimore, Md.: Penguin Books, 1957. Chapter 8, "The Psychology of Aesthetics," pp. 308-39.
- Fechner, G. T. *Vorschule der Aesthetik*. Theil 1 and 2. Leipzig: Breitkopf and Härtel, 1876.
- Goodman, Nelson. *Language of Art*. New York: Bobbs-Merrill, 1968.
- Husserl, Edmund. *Ideas: General Introduction to Pure Phenomenology*. (Translated by W. R. Boyce Gibson.) New York: Macmillan, 1931.
- Kaelin, E. F. *Method and Methodology in Literary Criticism*. *School Review* 72: 289-308; 1964.
- Kaelin, E. F. *An Existential-Phenomenological Account of Aesthetic Education*. Penn State Papers in Art Education No. 4. University Park: Penn. State Univ., 1968a.
- Kaelin, E. F. *Aesthetic Education: A Role for Aesthetics Proper*. *Journal of Aesthetic Education* 2: 51-66; 1968b.
- Kaelin, E. F. *Epoché and Relevance in Aesthetic Education*. *Aesthetic Concepts*. (Edited by R. A. Smith.) Urbana: Univ. of Ill. Press, 1970.
- Langer, Susanne K. *Philosophy in a New Key*. New York: Penguin Books, 1948.
- Langer, Susanne K. *Feeling and Form*. New York: Scribner's Sons, 1953.
- Langer, Susanne K. *Problems of Art*. New York: Scribner's Sons, 1957.
- Leijonhielm, Christer. *Colours, Forms and Art: Studies in Differential Aesthetic Psychology*. Stockholm: Almqvist and Wiksell, 1967.
- Margolis, Joseph. *The Language of Art and Art Criticism*. Detroit, Mich.: Wayne State Univ. Press, 1965.
- Merleau-Ponty, Maurice. *Le Langage Indirect et les Voix du Silence*. *Les Temps Modernes* 7: 2113-44; June 1952. 7: 70-94; July 1952.
- Morris, Charles W. *Science, Art and Technology*. *Kenyon Review* 1: 409-23; 1939a.
- Morris, Charles W. *Esthetics and the Theory of Signs*. *Journal of Unified Science* 7: 131-50; 1939b.
- Morris, Charles W. *Significance, Signification, and Painting*. *The Language of Value*. (Edited by R. Lepley.) New York: Columbia Univ. Press, 1957. Pp. 58-76.
- Munro, Thomas. *Toward Science in Aesthetics*. New York: Liberal Arts Press, 1956.
- Platt, John R. *Strong Inference*. *Science* 146: 347-52; 1964.
- Prall, David W. *Aesthetic Judgment*. New York: Crowell Apollo Editions, 1967. Chapters 3, 4, and 5.
- Pratt, Carroll C. *The Stability of Aesthetic Judgments*. *Journal of Aesthetics and Art Criticism* 15: 1-11; 1956.
- Santayana, George. *The Sense of Beauty*. New York: Scribner's Sons, 1896.
- Smith, Ralph A. *Patterns of Meaning in Aesthetic Education*. *Council for Research in Music Education* 5: 1-12; 1965.
- Smith, Ralph A. *Aesthetic Education at the University of Illinois*. *Journal of Aesthetic Education* 1: 123-26; 1966.
- Valentine, C. W. *The Experimental Psychology of Beauty*. London: Methuen, 1962.
- Villemain, Francis T. *Methodological Inquiry Into Aesthetic Subject Matter*. *Proceedings of the Seventeenth Annual Meeting of the Philosophy of Education Society*. (Edited by Robert Mason.) Syracuse, N.Y.: Philosophy of Education Society, Syracuse Univ., 1961.
- Weitz, Morris. *The Role of Theory in Aesthetics*. *Journal of Aesthetics and Art Criticism* 15: 27-35; 1956.
- Woodworth, Robert S. and Schlosbert, Harold. *Experimental Psychology*. Revised edition. New York: Henry Holt, 1954.

- Ziff, Paul. Art and the "Object of Art." *Essays in Aesthetics and Language*. (Edited by William Elton.) Oxford: Blackwell, 1954. Pp. 170-86.
- Ziff, Paul. *Semantic Analysis*. Ithaca, N.Y.: Cornell Univ. Press, 1960.

### Additional References

- Barron, Frank. Artistic Perception as a Factor in Personality Style. *Journal of Psychology* 33: 199-203; 1952.
- Beardsley, Monroe C. and Schueller, Herbert M. (editors). *Aesthetic Inquiry*. Belmont, Calif.: Dickenson, 1967.
- Beebe-Center, J. H. and Pratt, C. C. A Test of Birkhoff's Aesthetic Measure. *Journal of General Psychology* 17: 335-50; 1930.
- Beittel, Kenneth R. *Effect of Self-Reflective Training in Art on the Capacity for Creative Action*. Cooperative Research Report No. ED-016-900. University Park: Penn. State Univ., 1964.
- Birkhoff, G. D. *Aesthetic Measure*. Cambridge, Mass.: Harvard Univ. Press, 1932.
- Child, Irvin L. Personal Preferences as an Expression of Aesthetic Sensitivity. *Journal of Personality* 30: 496-512; 1962.
- Child, Irvin L. *The Problem of Objectivity in Esthetic Value*. Penn State Papers in Art Education, No. 1. University Park: Penn. State Univ., 1966.
- Davis, R. C. An Evaluation and Test of Birkhoff's Aesthetic Measure. *Journal of General Psychology* 15: 231-40; 1936.
- DeWare, H. A Comparison of Tests of Aesthetic Measure. *British Journal of Educational Psychology* 28: 29-49; 1933.
- Ecker, David W. *Research and Development Team for the Improvement of Teaching Art Appreciation in the Secondary School*. Cooperative Research Project No. ED-011-063. Columbus: Ohio State Univ., 1966.
- Eisner, Elliot W. *A Comparison of the Developmental Drawing Characteristics of Culturally Advantaged and Culturally Disadvantaged Children*. Cooperative Research Report No. ED-015-783. Stanford, Calif.: Stanford Univ., 1966a.
- Eisner, Elliot W. The Development of Information and Attitude Toward Art at the Secondary and College Levels. *Studies in Art Education* 8: 43-58; Autumn 1966b.
- Eisner, Elliot W. and Ecker, David W. (editors). *Readings in Art Education*. Waltham, Mass.: Blaisdell, 1966.
- Eysenck, H. J. The Empirical Determination of an Aesthetic Formula. *Psychological Review* 48: 83-92; 1941a.
- Eysenck, H. J. A Critical and Experimental Study of Color. *American Journal of Psychology* 54: 385-94; 1941b.
- Eysenck, H. J. "Type"-Factors in Aesthetic Judgment. *British Journal of Psychology* 31: 262-70; 1941c.
- Eysenck, H. J. General Factor in Aesthetic Judgments. *British Journal of Psychology* 31: 94-102; 1941d.
- Gallie, W. B. The Function of Philosophical Aesthetics. *Mind* NS 57: 302-21; 1948.
- Getzels, Jacob W. *Creative Thinking in Art Students: An Empirical Study of Problem-Solving in "Presented" and "Discovered" Situations*. Cooperative Research Report No. ED-003-079. Chicago: Univ. of Chicago, 1964.
- Hambridge, Jay. *Dynamic Symmetry*. Cambridge, Mass.: Harvard Univ. Press, 1926.
- Kaelin, E. F. *An Existentialist Aesthetic: The Theories of Sartre and Merleau-Ponty*. Madison: The Univ. of Wis. Press, 1962.
- Kaelin, E. F. Isness and Oughtness. *Art Education* 21: 6-10; Jan. 1968.
- Margolis, Joseph. Recent Work in Aesthetics. *American Philosophical Quarterly* 2: 182-92; 1965.
- McWhinnie, Harold J. A Review of Research on Aesthetic Measure and Perceptual Choice. *Studies in Art Education* 6: 34-41; Spring 1965.
- McWhinnie, Harold J. A Review of Research on Aesthetic Measure. *Acta Psychologica* 28: 363-75; 1968a.
- McWhinnie, Harold J. The Effects of a Learning Experience Upon the Preference for Complexity and Asymmetry in Fifth Grade Children. *California Journal of Educational Research* 10: 183-89; 1968b.

- Pepper, Stephen C. *The Basis of Criticism in the Arts*. Cambridge, Mass.: Harvard Univ. Press, 1956.
- Prall, D. W. *Aesthetic Analysis*. New York: Crowell Apollo Editions, 1967.
- Smith, Ralph A. (editor). *Aesthetics and Criticism in Art Education*. Chicago: Rand McNally, 1966.
- Stumbo, Hugh W. Three Bases for Research and Teaching in the Arts: Subjective, Objective, and Projective. *Studies In Art Education* 9: 21-30; Spring 1968.
- Wilson, Brent G. An Experimental Study Designed to Alter Fifth and Sixth Grade Students' Perception of Paintings. *Studies In Art Education* 8: 33-42; Autumn 1966.
- Ziff, Paul. Reasons in Art Criticism. *Philosophy and Education*. (Edited by Israel Scheffler.) Boston: Allyn and Bacon, 1958. Pp. 219-36.

## AUTHORS

- DAVID W. ECKER Address: New York University Title: Professor of Art Education Age: 38 Degrees: B.S., State Univ. College at Buffalo; M.S., Univ. of Wis.; Ed.D., Wayne State Univ. Specialization: Art Education; Aesthetic Education; Philosophy of Education.
- THOMAS, J. JOHNSON Address: Central Midwestern Regional Educational Laboratory, Inc. Title: Deputy Director Age: 40 Degrees: A.B., M.A., and Ph.D., Washington Univ. Specialization: Conceptual Learning and Measurement; Information Storage and Retrieval.
- EUGENE F. KAELIN Address: Florida State University Title: Professor of Philosophy Age: 43 Degrees: A.B. and M.A., Univ. of Mo.; D.E.S., Univ. of Bordeaux (France); Ph.D., Univ. of Ill. Specialization: Aesthetics; Phenomenology; Arts in Society.

## 4: ANTHROPOLOGICAL APPROACHES TO THE STUDY OF EDUCATION

PETER S. SINDELL\*

*McGill University*

George D. Spindler (1969) has delineated some of the fundamental characteristics of the anthropological method. Anthropologists are committed to ethnographic fieldwork entailing direct observation and participation in the flow of life surrounding them. Lengthy immersion in an alien culture (for at least a year initially, followed by return visits when possible), knowledge of the indigenous language and collection of a wide range of data utilizing diverse methods are generally considered to be requisites for gaining a sensitive and accurate understanding of a socio-cultural situation and its dynamics. To distinguish what is indiosyncratic, what is a function of the various social roles present, and what is attributable to subcultural or cultural patterning, anthropologists try to select informants who come from the different cultures and subcultures present and who occupy various statuses in the social system. When a group is experiencing rapid socio-cultural change (modernization, urbanization, revolution, etc.), anthropologists attempt to sample the spectrum of personal exposure to and participation in the change.

When gathering and analyzing data anthropologists pay constant attention to the interdependence of phenomena; they tend to think in terms of cultural patterns or configurations. They attempt to see how each discrete fact relates to the total matrix of other facts collected on the socio-cultural situation. Furthermore, there is a strong concern for the undeclared meanings and latent functions of behavior (the "covert culture") as well as for the immediately observable content and manifest functions. Despite this holistic orientation to what they study, anthropologists usually employ (implicitly or explicitly) a cross-cultural comparative framework when they present their findings and interpret them.

This anthropological perspective should encourage researchers to relate each school system, each school, and each classroom to the broader social and cultural contexts in which it exists. In relating a classroom or school to the national social system, for example, one can ask what cultural

---

\* Drs. Jacquetta Hill Burnett, University of Illinois, and George D. Spindler, Stanford University, served as consultants to Mr. Sindell on the preparation of this chapter. Mr. Sindell expresses his appreciation to the editors and authors who shared with him their unpublished and forthcoming material.



values and goals are reflected in the policy directives and curricular material which are issued by the national educational authorities, state or provincial school boards, and local school systems. To relate a school to the constituency it serves, one might ask what continuities and discontinuities the children experience in enculturation as they interact with parents, siblings, and kin or with schoolmates, teachers, and other school staff. With this analytic stance researchers also should seek to determine what cultural and social patterns influence each pupil, teacher, or administrator in his interactions with the others. One studies the culturally and socially conditioned values, attitudes, cognitive styles, self conceptions, role expectations, and modes of interpersonal interaction which pupils, teachers, and administrators bring with them into the school context from their various backgrounds. One also investigates the shared perceptions, values, and rituals from which student, staff, and school subcultures arise.

In reviewing anthropological studies of education and those which utilize anthropological methods, one discerns three principal foci of research: schools and their relations with the socio-cultural milieu in which they exist; the description and analysis of classroom processes; and the study of individual pupils and educators. The strengths and weaknesses of studies focusing on these three areas will be assessed briefly in this paper. Then some systematic methodological biases apparent in the literature will be discussed and some directions for future research will be suggested.

### *Schools and Their Socio-Cultural Milieux*

Methodologically this approach is closely related to that used in traditional anthropological community studies, but it emphasizes fieldwork in an educational setting as well as in the community at large. Investigators who conducted studies in this manner were concerned with education in rural Germany (Spindler, 1969; Warren, 1967), in Japan (Brameld, 1968; Singleton, 1967, 1968), in a Philippine barrio (Jocano, 1969), in a Canadian Indian (Kwakiutl) village (Wolcott, 1967, 1968a, 1969c), on Sioux (Wax, Wax, and Dumont, 1964) and Apache (Parmee, 1968) Indian reservations in the United States and in a residential school serving Canadian Indian children in the Yukon Territory (King, 1967).

The prime virtue of the community oriented approach is that it leads one to explore in detail the relationship between the school and its socio-cultural context: for example the intricate ways in which the child's familial socialization and peer group relations interact with his life in school. Wolcott's Kwakiutl data demonstrate dramatically how teachers with cultural orientations which conflict with those of their pupils and the community are stymied in their educative efforts (Wolcott, 1967, 1968a, 1969c). In contrast, George and Louise Spindler's current research on

urbanization and education in a rural German village (Spindler, 1969) shows how a school can function to provide children from varying backgrounds with the knowledge to cope effectively with an emerging urban culture.

Wolcott, Spindler and the others who have chosen this broad focus view the school as only one of the enculturative agencies that affects the child. Consequently the ways in which the school reinforces or conflicts with the child's enculturative experiences in other milieux is one important question which is posed by investigators. Sindell's (1968) observations of six year old Cree Indian children from northern Quebec in a residential school, in the bush with their parents, and at a trading post clarify how conflicts in cultural values, role expectations, and styles of interpersonal interaction affect school adaptation and academic learning. (cf. Rohner, 1965.) For example because the Cree definition of aggressive behavior is different from the Euro-Canadians', the Cree students frequently thought the teacher was angry when he raised his voice to emphasize a point.

Burger has worked extensively in this area. He prepared a manual (Burger, 1968) for teachers in the Southwestern United States which specifies precisely some of the cultural differences between Anglo, Hispano-American, and Amerindian groups and suggests ways teachers can educate children using methods and curricular materials which are culturally appropriate.

If one thinks of the school itself as a community in interaction with other communities, as Singleton does (1967, 1968), then the influence of the societal macrosystem on the local school is revealed. Singleton analyzed the Nichu school as a group of people with common interests and loyalties which interacts with other communities: (1) the local socio-geographic community of persons residing in the school district, (2) the administrative community of persons and organizations having control over and responsibility for the school, and (3) the union-professional community of Japanese teachers. By emphasizing the school's relationships with other institutions in Japanese society this conceptual model lends a sharp vertical dimension to Singleton's description and analysis which is lacking in most other studies.

The virtue of such broad, holistic studies is their emphasis on the many ways in which teaching and learning in a school are affected by the social and cultural processes occurring in the surrounding milieu, ranging from familial socialization to urbanization and modernization. The problem with these studies is that they have too often stayed at the broadly descriptive level. Although one major aim of such empirical case studies is to produce interpretive generalizations through induction (Spindler, 1967), the data are rarely detailed enough in any one problem area to allow confident generalizations. One reason for this problem is that the task under-

taken is too large for one ethnographer to accomplish well even in a year (especially if the ethnographer has teaching responsibilities). Perhaps if a research team were used instead of a single ethnographer the configurative understanding which comes with this approach could be retained but the empirical precision of the results could be increased. Intensive studies of traditional enculturation such as those produced in the Six Cultures Study (Whiting, 1963), by Read (1968) and by Williams (1969) should be articulated with studies of education.

The "Case Studies in Education and Culture," (Jocano, 1969; King, 1967; Singleton, 1967; Warren, 1967; and Wolcott, 1967) furnish a great deal of comparative material which has not been utilized either by educators or anthropologists in developing cross-culturally valid interpretive generalizations or testable hypotheses.

A final criticism of these holistic studies of schools and their sociocultural milieux lies in their bias toward the teacher's view of reality. Children were described in classrooms and familial socialization was usually discussed briefly but only Wax, Wax, and Dumont (1964) and Spindler (1969) intentionally interacted extensively with children in a variety of social situations and utilized them as key informants, informally and through structured interviews.

### *Classroom Processes*

In the second major research strategy, researchers collect and analyze data about classroom processes *per se*. With the exception of Cohen's work (forthcoming), these studies are strictly circumscribed in their focus rather than holistic and far ranging. Investigators using this approach have either studied one classroom intensively or examined an extensive sample of classrooms.

Reading Smith and Geoffrey (1968) is an exciting experience for an anthropologist because they have explicitly used anthropological techniques to generate verifiable hypotheses as well as to build an ethnographic description of one urban classroom (based upon a full semester of fieldwork). But reading Smith and Geoffrey is also very frustrating because they rarely related the behavior they observed to anything outside the classroom or school. They asked almost purely psychological questions of the data. The teacher and pupils seemed to exist in a cultural and social vacuum. One has to search very hard even to discover the pupils' ages, ethnic, and racial backgrounds. Pupil beliefs and faculty norms were postulated in some of Smith and Geoffrey's models but these beliefs and norms were only described in a cursory manner. No data were given on pupils' attitudes or values, pupils' cultural and social backgrounds, the teacher's background or values, or the neighborhood in which the school functioned. Leacock



(in press) underscored the importance of data such as these when she stated:

thus the general character of a school—its culture so to speak—follows in great part from the neighborhood it serves, and this general atmosphere is more relevant to what children are learning than the personality of the individual teacher.

Despite my criticisms Smith and his colleagues have contributed greatly to the development of new methods in educational ethnography. (See Smith, 1967, 1969; Smith and Geoffrey, 1968; Smith and Pohland, 1969.) In the first study the ethnographer (Smith) and the teacher (Geoffrey) collaborated totally in collecting and analyzing data and generating theoretical propositions about it. During the long period when they studied the complexities of an urban classroom Geoffrey wrote fieldnotes after teaching every day and Smith wrote running process records during classtime (he was present 80 percent of the time) and then dictated more general summary observations and interpretations after school. It is clear from Smith and Geoffrey's findings that the intensive study of one classroom situation over a long period of time is invaluable. For example, only in this way could they have discovered how particular patterns of teacher-pupil interaction develop over time, such as Geoffrey's unspoken agreement with certain difficult students not to demand academic work from them if they did not disrupt the class.

It is clear from Smith and Geoffrey's study and from Wolcott's experience (1969b) that a teacher working alone is limited in his ability to do ethnographic work because he is emotionally involved in the very processes he is studying and because of the tremendous amount of social interaction he must initiate and react to while teaching. In addition, being a teacher or being perceived as a teacher limits his access to a wide range of data, particularly about pupils and parents' attitudes and perceptions (Singleton, 1968). But as Smith and Geoffrey (1968) pointed out, frequently only the teacher knows why he did something or what factors influenced his behavior. This point has general significance and, in my opinion, Smith and Geoffrey's innovative use of a teacher-ethnographer team is a technique that educational ethnographers should adopt whenever feasible. Another valuable method employed by Smith and Geoffrey was their use of "interpretive asides." While he was writing field notes Smith took time out to write down any insights or tentative interpretations which occurred to him as he was observing. These insights later played a major role in developing hypotheses. The ways in which Smith and Geoffrey developed hypotheses, returned to the ethnographic data to test them, and then elaborated them are instructive to anyone working with this kind of data.



Since they conducted their initial study Smith and his colleagues have utilized microethnographic methods in analyses of an apprentice teaching program (Smith, 1968; Connor and Smith, 1967), an innovative elementary school (Smith and Keith, 1967), and a computer assisted learning program (Smith and Pohland, 1969). In addition, verification studies of hypotheses developed by Smith and Geoffrey have begun, for example on the relationship between "teacher awareness" and pupil esteem (Smith and Kleine, in press).

In contrast to Smith and Geoffrey's intensive focus, several investigators have chosen to observe in a sample of classrooms. Leacock (1969) and her associates at the Bank Street College of Education studied and contrasted two classrooms in each of four schools serving pupils from low income Negro, middle income Negro, low income White and middle income White populations. Data on each classroom was gathered from three periods of observation (by two observers simultaneously), two teacher interviews, and a brief pupil questionnaire.

Project TRUE (Eddy, 1967; Moore, 1967), like the Bank Street College Study, relied primarily on brief observations in a series of urban classrooms and interviews with school personnel. The major disadvantages of this approach are that adequate rapport with pupils and staff cannot possibly develop when such a brief period is spent in each setting and that the potentially disruptive effect of the observer is maximized. Furthermore, as with Smith and Geoffrey, these studies did not involve ethnographic fieldwork in the communities served by the schools so that the influence of the pupils' socio-cultural background was largely ignored.

Cohen (forthcoming) also utilized observations in a series of classrooms when he set out to analyze the educational systems of Israel and the United States, but his analytic focus was on the national social systems and how they influenced the educational process. Cohen contended that individuals live less and less in communities and more and more within nationally anchored institutions. In two brilliant papers, (1969, in press) he argued from an evolutionary standpoint: (1) that the development of schools is a characteristic feature of state societies; (2) that state societies need to subvert local sources of solidarity, loyalty, and authority in favour of universalistic orientations; and (3) that state societies need to legitimate their authority by creating a set of uniform ideological symbols for their future participants. Education, in Cohen's view, is of central importance in fulfilling these functions in state societies.

Cohen felt, correctly in my opinion, that anthropologists have generally neglected national social systems and have failed to develop methods which are appropriate to the hierarchical complexities of modern nations. To date this criticism is also relevant to the anthropological study of education

since only Singleton's (1967) study of Nichu displayed concern for these issues.

Assuming that "each of a nation's centrally controlling bureaucracies and its associated sphere of activity is a microcosm of the whole, synchronically and diachronically," Cohen attempted to document and analyze the bureaucratic style of the centralized bureaucracies and the contents of the policies they promulgated. Cohen's method appears to promise greater understanding of the national institutional forces which affect the school.

In sum, both approaches to the study of classroom processes, intensive observation in one classroom and brief observation in several classrooms, are of limited value unless accompanied by fieldwork which relates the phenomena under study either to the local social and cultural situation or to the national social system.

### *Individual Pupils and Educators*

An exciting new approach which combines the strengths of the community oriented studies with Smith and Geoffrey's microethnographic procedures is presently being used and developed by Jacquetta Burnett (1968). Her theoretical interest lies in discovering the ways in which differences between household culture and school culture in an urban Puerto Rican area in Chicago affect schooling. For example, how do Puerto Rican notions of "machismo" affect the interaction of male students with their female teachers. She began her study by taking a "sample of egocentric networks," i.e., she chose thirty seventh graders from one school and followed each one for several months through the network of social contacts in which he was involved. Through interviews and extensive observations this method reveals how youths and adults view each other and how they view common problems. Furthermore, it assists the analyst in discovering the significant cultural differences which are affecting the pupil in his interaction with peers, teachers, and kin. To control what behavior is related to cultural differences and what is a function of poverty and urban life style, random samples were selected of: five girls and five boys who had attended school in Puerto Rico for five years or more, five girls and five boys of Puerto Rican descent who had had all of their schooling in Chicago, and five girls and five boys of non-Puerto Rican descent.

To compare "observed events" occurring in the classroom and in the community, Burnett and her team have been developing ways to describe events in terms of their location in space, the physical objects and people present, the actions and interactions (verbal and non-verbal) occurring, and the order in which these various elements appear. Following Harris (1964) they are developing "actor types", "placespace types", "object types",

and "absolute and relative time" as "stage coordinates". It will not be possible to make valid cross-cultural comparisons and generalizations unless precise ways to describe the educational phenomena observed are developed. As well as describing segments of the behavior stream (Barker, 1963) which seem to have some internal integrity ("events"), Burnett and her co-workers are gathering data on the participants' values, desires, and cognitive maps by asking informants why particular activities are occurring.

Wolcott's (1968b, 1969b) aim is different from Burnett's. Wolcott sought to describe and analyze the behavior of an elementary school principal and those who interact with him—teachers, pupils, spouse, family; but his essential method was very similar to egocentric network analysis. His point of departure, like Burnett's, was an individual and the whole range of his social contacts. In his two year study, Wolcott (1968b, 1969b) collected three kinds of material: enumerative and census data; protocols and field-notes based upon participant observation; and interviews with informants, including the key informant, the principal. In addition to the usual documentary materials—notes, records, etc.—Wolcott mapped and photographed the school and the neighborhood and recorded "time and motion data" on the principal's behavior over a two week period. In his discussion of participant observation, Wolcott stressed the importance of keeping up with one's notes; he stopped observing until he had written up all prior observations. Smith and Geoffrey (1968) and Wolcott (1969b) both pointed out that new events obscure older ones. Smith and Geoffrey noted that even talking about some event before writing it down tends to confuse its significance and obscure its details.

Wolcott emphasized the importance of rapport in interviewing informants; he stated that he did not conduct interviews with the teachers in the school until he had been there six months. As in Smith and Geoffrey's collaboration, Wolcott had the willing involvement of the principal in the data collection phase of the research in terms of recounting events of significance, sharing reflections and feelings, allowing formal interviews, etc., but unlike Geoffrey the principal was not involved actively in the data analysis. In another context Wolcott (in press) warned anthropologists of the need to display the same careful attention to building rapport and doing careful fieldwork in educational settings as they do in more exotic situations.

Warren (1968) is presently engaged in research on the teaching experience which is similar to that of Burnett and Wolcott in emphasizing observation of particular teachers in all of the settings, classroom and extra-classroom, in which they play their roles. The teacher-encounter, "a natural appearing unit of interpersonal interaction which constitutes for the teacher a recognizable element in her occupational world", is comparable to Burnett's "event" and plays a similar role in comparisons of teacher's classroom experiences and her (in this case school related) experiences outside the



classroom. Warren postulated "physical setting", "population", and "activity" as three elements in the teacher encounter. As in Burnett's study if the parameters of teacher behavior can be isolated and identified, they can be used in testing hypotheses cross-culturally.

## Conclusions

A review of the literature clearly shows that anthropological approaches to the study of education concurrently need to be (1) broad enough to encompass the full range of socio-cultural influences which affect the schools, pupils, and educators; (2) microscopic enough in their focus to result in precise descriptions of the phenomena under study; and (3) theoretically oriented enough to generate hypotheses about the interrelationships of the data discovered. In my opinion the most promising approach would be the study of individually oriented networks coupled with lengthy fieldwork, microethnographic methodology, interviews and time sampling. It is clear, however, that anthropologists have a long way to go before they have adequate ways to describe "real" events in ways which allow for precise multidimensional comparisons. Whatever methods are used in the future, it is important to correct certain biases which seem to characterize the field at present.

First, there seems to be a bias toward studying reality only as adults see it. ". . . schools are adult institutions in which children participate; they are also childhood institutions in which adults play all important roles" (Cohen, 1969). Most studies reflect the first perspective expressed in the quotation. Consequently when studying schools, anthropologists rarely interview students in depth about their feelings, attitudes and values. Nor do they usually do participant observation with children outside the classroom. One difficulty in this area is the difference between researchers and children in age and status. However, one could utilize pupils directly as research assistants and use college students or graduate students in this kind of research, especially when studying with adolescents.

Anthropologists devote a disproportionate amount of their research effort to the elementary and junior high school grades while neglecting high school and university students as subjects for ethnographic research. Burnett's study of a rural midwestern high school (1969, in press) will help to fill this gap on the high school level. However, Bushnell's ethnographic sketch of Vassar (1962) and *Boys in White* (Becker et al., 1961) are among the few ethnographically oriented studies of student subcultures in universities.

A final bias one can discern is an overemphasis on observational methods and a neglect of cognitive approaches. While reading ethnographic studies of education one is constantly struck especially in the urban studies,



by the teacher's rich vocabulary of terms for students, "troublemaker", "discipline case" and "joker" for example. One way to discover teachers' attitudes toward pupils would be to explore the criteria which teachers use in labeling pupils. These perceptions clearly condition teacher-pupil interaction and affect a student's "reputation" throughout his school career. Recently anthropologists have been developing formal ethnographic procedures to elicit informants' perceptions of their social world. Frake's paper on the ethnographic study of cognitive systems, Black and Metzger's study of law, and Goodenough's rethinking of status and role (all in Tyler's, 1969, book of readings, *Cognitive Anthropology*) present methods which are relevant to the analysis of a problem such as that noted above.

Gay and Cole's experimental studies of Kpelle cognition (1967) represent another line of inquiry which deserves development, i.e. experimental determination of cultural differences in cognitive style and how these affect education. Gay and Cole also alert us to the importance of learning modes or meta-learning (cf. Wolcott, 1969a).

Because of their observational bias anthropologists also have neglected the effect of education on the development of personal and cultural identity. Goodenough (1963) has discussed this problem in theoretical terms but little empirical work has been done in this area. The fruitfulness of this approach is indicated in the research that Wintrob and Sindell (1968) have done on education and identity conflict among Cree Indian adolescents from northern Quebec. All adolescents from two Cree bands attending school in 1967-1968 and a control group of adolescents who had had little or no educational experience were interviewed. The Adolescent Adjustment Interview (AAI), consisting of some one hundred items, was designed to investigate each student's educational, occupational and social aspirations, the degree of parental opposition to these aspirations, the extent of the individual's confidence in achieving his goals along each of the three parameters, and the prominence of his feelings of inadequacy, anxiety, and depression. Ethnographic fieldwork in the traditional milieu of trading post and hunting-trapping group and in the urban milieu where the students attended school provided essential information on parental attitudes toward education, family interaction patterns, cultural values and role expectations, and students' experiences in the urban setting. Collating data gathered on individuals and families through clinical interviews and the AAI and from ethnographic studies in the community and school elucidated the nature of the students' identity conflict and its relationship to students' educational experiences.

In closing it is appropriate to call for more empirical discussion of methods and methodology in educational anthropology. Too many authors devote scant attention to the methods they have utilized in obtaining and analyzing their data. Too many publications do not include the texts of

interviews or questionnaires administered and do not specify precisely where, when, and under what conditions the research was done. Only if we discuss our successes and failures and compare them can we develop methods which have general utility and cross-cultural replicability.

## Bibliography

- Barker, Roger G. (editor). *The Stream of Behavior. Explorations of its Structure and Content*. New York: Appleton-Century-Crofts, 1963.
- Becker, Howard S. et al. *Boys in White: Student Culture in Medical School*. Chicago: Univ. of Chicago Press, 1961.
- Brameld, Theodore. *Japan: Culture, Education, and Change in Two Communities*. New York: Holt, Rinehart and Winston, 1968.
- Burger, Henry G. "Ethno-pedagogy": *A Manual in Cultural Sensitivity, with Techniques for Improving Cross-Cultural Teaching by Fitting Ethnic Patterns*. Albuquerque, N.M.: Southwestern Cooperative Educational Laboratory (117 Richmond Drive, N.E.), 1968.
- Burnett, Jacquetta Hill. *Event Analyses in the Microethnography of Urban Classrooms*. Paper presented to the Annual Meeting of the American Anthropological Assoc., Nov. 1968. Urbana: Bureau of Educational Research, Univ. of Ill. (Mimeo.)
- Burnett, Jacquetta Hill. Ceremony, Rites and Economy in the Student System of an American High School. *Human Organization* 28: 1-10; 1969.
- Burnett, Jacquetta Hill. *Pattern and Process in Student Life: A Study of Custom and Social Relationship among the Students in an American High School*. New York: Teachers College Press, in press.
- Bushnell, John H. Student Culture at Vassar. *A Psychological and Social Interpretation of the Higher Learning*. (Edited by Nevitt Sanford.) New York: John Wiley and Sons, 1962. Pp. 489-514.
- Cohen, Yehudi A. Schools and Civilizational States. *The Social Sciences and the Comparative Study of Educational Systems*. (Edited by Joseph Fischer.) New York: International Textbook, 1969. Pp. 39-105.
- Cohen, Yehudi A. The Shaping of Men's Minds. *Adaptations to Imperatives of Culture. Anthropological Perspectives on Education*. (Edited by Murray Wax and Stanley Diamond.) New York: Basic Books, in press.
- Cohen, Yehudi A. *A Method for the Anthropological Study of National Social Systems*, forthcoming.
- Connor, William H. and Smith, Louis M. *Analysis of Patterns of Student Teaching*. Final Report 5-8204. Washington, D.C.: U.S. Office of Education, Bureau of Research, 1967.
- Eddy, Elizabeth M. *Walk the White Line. A Profile of Urban Education*. Garden City, N.Y.: Doubleday, 1967.
- Gay, John and Cole, Michael. *The New Mathematics and an Old Culture. A Study of Learning Among the Kpelle of Liberia*. Case Studies in Education and Culture. New York: Holt, Rinehart and Winston, 1967.
- Goodenough, Ward H. *Cooperation in Change. An Anthropological Approach to Community Development*. New York: Russell Sage Foundation, 1963.
- Harris, Marvin. *The Nature of Cultural Things*. New York: Random House, 1964.
- Jocano, F. Landa. *Malibog: Education in a Philippine Barrio*. Case Studies in Education and Culture. New York: Holt, Rinehart and Winston, 1969.
- King, A. Richard. *The School at Mopass. A Problem of Identity*. Case Studies in Education and Culture. New York: Holt, Rinehart and Winston, 1967.
- Leacock, Eleanor Burke. *Teaching and Learning in City Schools: A Comparative Study*. New York: Basic Books, 1969.
- Leacock, Eleanor Burke. Theoretical and Methodological Problems in the Study of Schools. *Anthropological Perspectives on Education*. (Edited by Murray Wax and Stanley Diamond.) New York: Basic Books, in press.

- Moore, G. Alexander. *Realities in the Urban Classroom. Observations in Elementary Schools*. Garden City, N.Y.: Doubleday, 1967.
- Parmee, Edward A. *Formal Education and Culture Change. A Modern Apache Indian Community and Government Education Programs*. Tucson: Univ. of Ariz. Press, 1968.
- Read, Margaret. *Children of Their Fathers. Growing Up Among the Ngoni of Malawi*. Case Studies in Education and Culture. New York: Holt, Rinehart and Winston, 1968.
- Rohner, Ronald P. Factors Influencing the Academic Performance of Kwakiutl Children in Canada. *Comparative Education Review* 9: 331-40; 1965.
- Sindell, Peter S. Some Discontinuities in the Enculturation of Mistassini Cree Children. *Conflict in Culture. Problems of Developmental Change Among the Cree*. (Edited by Norman A. Chance.) Ottawa: Canadian Research Centre for Anthropology, 1968. Pp. 83-92.
- Singleton, John. *The Ethnography of a Japanese School: Anthropological Field Techniques and Models in the Study of a Complex Organization*. Paper presented at the Annual Meeting of the American Anthropological Assoc., Nov. 1968. Pittsburgh, Pa.: International and Development Education Program, Univ. of Pittsburgh. (Mimeo.)
- Smith, Louis M. The Microethnography of the Classroom. *Psychology in the Schools* 4: 216-21; 1967.
- Smith, Louis M. *Classroom Social Systems in Teacher Education*. Central Midwestern Regional Educational Laboratory, Occasional Paper No. 4. St. Ann, Mo.: the Laboratory (10646 St. Charles Rock Road), 1968. (Mimeo.)
- Smith, Louis M. and Geoffrey, William. *The Complexities of an Urban Classroom. An Analysis toward a General Theory of Teaching*. New York: Holt, Rinehart and Winston, 1968.
- Smith, Louis M. and Keith, Pat M. *Social Psychological Aspects of School Building Design*. Final Report No. 3-223. Washington, D.C.: U.S. Office of Education, Bureau of Research, 1967.
- Smith, Louis M. and Kleine, Paul F. *Teacher Awareness: Social Cognition in the Classroom. The School Review*, in press.
- Smith, Louis M. and Pohland, Paul A. *Grounded Theory and Educational Ethnography: A Methodological Analysis and Critique*. St. Ann, Mo.: Central Midwestern Regional Education Laboratory, 1969. (Mimeo.)
- Spindler, George D. (editor). *The Character Structure of Anthropology. Education and Culture*. New York: Holt, Rinehart and Winston, 1967.
- Spindler, George D. Foreword. *The New Mathematics and an Old Culture. A Study of Learning among the Kpelle of Liberia*. New York: Holt, Rinehart and Winston, 1967.
- Spindler, George D. *Urbanization and Education in a Rural German Village*. Stanford, Calif.: Department of Anthropology, Stanford Univ., 1969. (Mimeo.)
- Tyler, Stephen A. (editor). *Cognitive Anthropology*. New York: Holt, Rinehart and Winston, 1969.
- Warren, Richard L. *Education in Rebhausen. A German Village. Case Studies in Education and Culture*. New York: Holt, Rinehart and Winston, 1967.
- Warren, Richard L. *Teacher Encounters: A Typology for Ethnographic Research on the Teaching Experience*. Paper presented to the Annual Meeting of the American Anthropological Assoc., Nov. 1968. Stanford, Calif.: Stanford Center for Research and Development in Teaching. (Xerox.)
- Wax, Murray L.; Wax, Rosalie H.; and Dumont, Robert V., Jr. *Formal Education in an American Indian Community*. SSSP Monographs: A Supplement to Social Problems, Vol. 2, No. 4; 1964.
- Whiting, Beatrice B. (editor). *Six Cultures. Studies of Childrearing*. New York: John Wiley and Sons, 1963.
- Williams, Thomas Rhys. *Dusun Enculturation. Case Studies in Education and Culture*. New York: Holt, Rinehart and Winston, 1969.
- Wintrob, Ronald M. and Sindell, Peter S. *Education and Identity Conflict among Cree Indian Youth. A Preliminary Report*. Annex 3 of the Final Report, McGill Cree

- Project. Ottawa: Rural Development Branch, Dept. of Regional Economic Expansion, 1968.
- Wolcott, Harry F. *A Kwakiutl Village and School*. Case Studies in Education and Culture. New York: Holt, Rinehart and Winston, 1967.
- Wolcott, Harry F. *Classroom Learning—Kwakiutl Style*. Paper presented at the American Educational Research Assoc., Feb. 1968a. Eugene: Center for Advanced Study of Educational Administration, Univ. of Ore. (Mimeo.)
- Wolcott, Harry F. *An Ethnographic Approach to Studying School Administration*. Paper presented to the Annual Meeting of the American Anthropological Assoc., Nov. 1968b. Eugene: Center for Advanced Study of Educational Administration, Univ. of Ore. (Mimeo.)
- Wolcott, Harry F. *Concomitant Learning: An Anthropological Perspective on the Utilization of Media*. *Educational Media: Theory into Practice*. (Edited by Raymond L. Winan and Wesley C. Meierhenry.) Columbus, Ohio: Charles B. Merrill, 1969a, Pp. 217-47.
- Wolcott, Harry F. *The Elementary School Principal: Notes from a Field Study*. Eugene: Center for Advanced Study of Educational Administration, Univ. of Ore., 1969b. (Mimeo.)
- Wolcott, Harry F. *The Teacher as an Enemy*. Eugene: Center for Advanced Study of Educational Administration, Univ. of Ore., 1969c. (Mimeo.)
- Wolcott, Harry F. *Handle with Care: Necessary Precautions in the Anthropology of Schools*. *Anthropological Perspectives on Education*. (Edited by Murray Wax and Stanley Diamond.) New York: Basic Books, in press.

## AUTHOR

PETER S. SINDELL Address: McGill University (Canada) Title: Assistant Professor of Anthropology Age: 27 Degrees: A.B., Harvard Univ.; M.A., Stanford Univ.; Ph.D. (candidate), Stanford Univ. Specialization: Educational Anthropology.





# The Journal of Teacher Education

*Published quarterly by The National  
Commission on Teacher  
Education and Professional Standards  
National Education Association  
1201 Sixteenth Street, N.W.  
Washington, D.C. 20036*

The only national journal devoted entirely to the education of teachers.

- Research
- Book reviews
- Critical analyses of teacher education
- Editorial comments
- A summary of new experiments, surveys, and programs

Highlights—Winter 1969

Symposium—Teacher Education: White Racism and  
Inner City Schools  
*Arthur Pearl, Editor*

Modern Teacher Education: Some Reflections on Its  
Ideological Base  
*Stanley R. Moore*

Technological Forces and the Teacher's Changing Role  
*Walt Le Baron*

A special book review section containing three reviews of  
*Teachers for the Real World*, one of the more important publi-  
cations of the National Institute for Advanced Study in  
Teaching Disadvantaged Youth

With other interesting articles, plus the regular departments:  
With the Researchers  
Of Special Significance  
Browsing Through the Bookshelves

Subscription rates: single issue, \$2.50; one year, \$8.00; three years, \$21.00

☐ Check enclosed ☐ Order enclosed ☐ Bill me

Name .....

Company .....

Address .....

City ..... State ..... Zip Code .....

## 5: PHILOSOPHY, EDUCATION AND THE DOOMSDAY THREAT

LEONARD JOSEPH WAKS\*

*Stanford University*

A decade ago, academic philosophy of education was at a standstill. The gap that separated educational philosophers in schools of education from professors in departments of philosophy made it difficult for new developments in philosophical method and technique to have impact on the philosophy of education.

In the past decade, the gap has narrowed. In aim and method, philosophy of education is now in greater harmony with general philosophical currents. (See Scheffler, 1969, 1965; Peters, 1966; Wilson, Williams and Sugerman, 1967. Original articles are collected in Archambault, 1965, and Peters, 1967. A collection of important reprints is found in MacMillan and Nelson, 1968.)

The dominant philosophical tendency in the English speaking world since the beginning of this century has been the rich and varied movement known as "analytic philosophy" or "conceptual analysis." (For aims and achievements of analytic philosophy see Urmson, 1956; Passmore, 1967; Warnock, 1958; and White, 1956. A collection of influential papers exhibiting an analytic approach to philosophy is found in Flew's book, 1965. A clear explanation of the techniques of analytic philosophy appears in Gorovitz and Williams's work, 1966.) To an extent greater than before, this movement is making its influence felt on philosophical writing on education. Hence, an examination of recent developments in "research method and technique" in the philosophy of education must concentrate on the influence of analytic philosophy.

Unfortunately it is not possible to pin this movement down to a few central theses. During the first two thirds of the century several tendencies developed within the analytic philosophy movement. These different strands within analytic philosophy sometimes differ on fundamental points. Rather than having a set of central theses in common, these conflicting positions exhibit certain family resemblances.

---

\*Drs. Hugh G. Petrie, Northwestern University, and Peter J. Goldstone, University of Maryland, served as consultants to Dr. Waks in the preparation of this chapter. F. Halfdan Pratt and Robin Hull Webber helped in the preparation of the manuscript and bibliography.

One purpose of this chapter is to identify some of these family traits and to illustrate them by providing some specimen analyses. Another purpose is to consider the value of the methods and techniques for the philosophy of education. On the basis of this discussion, I offer recommendations concerning the direction which research in the philosophy of education and the organization of research training in philosophy of education should take in the next decade, should there be a next decade.

### *Features of Analytic Philosophy*

William James (1955, p. 46) wrote that pragmatism "does not stand for any special results. It is a method only . . . . It has no dogmas, and no doctrine save its method . . . . the pragmatic method is an attitude of orientation . . . . The attitude of looking away from first things, principles, 'categories,' supposed necessities; and of looking towards last things, fruits, consequences, fact." Like pragmatism, with which it has overlapped to some extent, the analytical philosophy movement can be better characterized by its "method" or "attitude of orientation" than by any special doctrines.

Above all, analytic philosophy is concerned with language. Analytic philosophers believe that at least one important reason for the existence of philosophical problems is a defect in either the languages which are employed or the appreciation of the way in which these languages work. Consequently, analytical philosophers believe that a greater understanding of language will be of crucial importance for solving (or as some would rather say, dissolving) philosophical problems.

This concern with language does not originate with analytic philosophy. Plato, Locke, and Hume, among others, thought that matters of language and meaning were of central philosophical importance. There are some differences in approach, however, which set analytic philosophers apart from these classical philosophers. First, none of these earlier theorists placed quite the emphasis on language that most philosophers writing in English would today. Second, the forms which the interest in language takes are somewhat novel.

Throughout the development of the movement there were some among the analytic philosophers who concentrated on the construction of special, "artificial" languages within which philosophical problems could be solved (or in which they could not even be stated, although the languages were adequate for other purposes). The idea of constructing a perfect language within which intellectual problems could be solved dates back at least to Leibnitz. Yet it was not until the present century that this "ideal language" program received a significant share of attention.

Some philosophers within the analysis school have believed that this program is fundamentally mistaken and that there could never be such

an "ideal" language. Among these, some have discouraged even the piecemeal use in philosophy of specially constructed artificial languages; they have argued that there is nothing "wrong" with the natural languages of everyday and scientific use. These philosophers have attempted to understand the subtleties of ordinary language, believing that failure to understand how ordinary language actually functions is responsible for a variety of intellectual traps.

A substantial number of these philosophers have been attracted by the "ordinary language" philosophical *program*. These "ordinary language philosophers" believe that *all* philosophical problems arise because of these traps, and that by investigating ordinary language with great care, these traps can be detected and philosophical problems laid to rest. The increased attention to and respect for the details of ordinary language which their investigations have stimulated are, I suspect, of permanent philosophical importance.

The ideal-language philosophers had little direct influence on the philosophy of education, but in recent years the ordinary language approach has been very influential. The ordinary language philosophy *program* has not attracted many educational philosophers. On the contrary, even those philosophers of education who have been most willing to utilize ordinary language analysis have shown little interest in the *program* of reducing philosophy to analysis. Indeed, the program of ordinary language philosophy, which has been taken to discredit the more traditional tasks of the philosophy of education by denying them the label "philosophy," has been an obstacle to the use of language analysis in education. Many of the more traditional educational philosophers have felt that adopting the methods of analytic philosophy committed them to adopting the philosophical program as well, with a consequent commitment to alter their aims. This is discussed in the next section.

The emphasis on the importance of language is followed by a related attitude towards the philosophical enterprise. Even those analytic philosophers who have been chary of making any commitment to a philosophical program have generally believed the discipline of philosophy to be more narrowly restricted than either non-philosophers or philosophers of other persuasions. Most analytic philosophers have felt that "deep" and "unanswerable" questions, questions about how to live one's life, questions like "Should I commit suicide?" are in no way within their special province. Even when dealing with questions that have a clear and valid traditional claim to the label "philosophical," these philosophers are apt to take a somewhat unusual attitude.

It is easy to understand why these philosophers refuse to accept responsibility for telling people whether or not to commit suicide or how to run their lives. The intellectual satisfaction that philosophy affords



most often accompanies philosophical thinking for which the criteria of success are fairly clear. Academic philosophers of all schools are unwilling to spend precious time answering the unanswerable for the straightforward reason that one clearly can not. Nonetheless, analytic philosophers are even more concerned with the "tough-mindedness" of their approach than other philosophers. In practice, this tough-mindedness means excluding any questions which can not be handled securely.

One important result of this limitation is that analytic philosophers are likely to consider the methodological objectivity of philosophy to be on an equal footing with that of other areas of cognitive knowledge. Another important result is that many problems and questions which do not yield to the favored techniques of philosophers are neglected regardless of importance.

When dealing with traditional philosophical problems, the analytic philosopher is likely to believe that even if his solution or resolution of these problems is on an equal footing with the solution of a scientific problem, the problems and the traditional theories which have been proposed to solve the problems are not on an equal footing with scientific problems. Indeed, the analytic philosopher often regards the problems of philosophy and the traditional philosophical theories as confused and misguided, if not utter nonsense.

As I indicated above, the analytic philosopher may believe that inattention to the details of language is responsible for the generation of these problems. Similarly, he may believe that the philosophical theories designed to solve (rather than dissolve) the problems are misguided simply because they take the problems at face value.

For example, in answer to some of the "main problems of philosophy" many of the more traditional philosophers have proposed special philosophical theories like these: that no one ever knows anything about the world at all; that nothing ever really happens before anything else; that the only thing which is worthwhile at all is a sensation of pleasure which is very much like an itch. When these philosophical theories have compromised more ordinary concepts, some traditional philosophers have asserted that the ordinary way of thinking about these matters is confused and incorrect.

Philosophers of the analytic persuasion have been very skeptical about the existence of special philosophical knowledge. In particular, those influenced by ordinary language analysis are skeptical about the claim that ordinary scientific and everyday concepts are confused. Instead of building philosophical theories or "Isms" these philosophers often restrict themselves to an underlaborer's role. For example, rather than issuing a theory about the value of anything, these philosophers attempt to give a coherent account of the concepts and expressions employed in value judgments made

by those concerned and competent to make them. At first glance, this underlaborer conception of philosophy does not appear unreasonable, since it is not professors of philosophy but horsemen who are especially qualified to evaluate horses. Whether or not there is any special philosophical knowledge is considered in the next section.

For analytic philosophers, then, one common conception of the philosopher, the conception of a man with a body of doctrines on deep questions of life, is thoroughly rejected. Most of these philosophers regard philosophy as an intellectual activity consisting of description and clarification of expressions in scientific and everyday discourse, and not as a body of doctrines. Philosophy, it is said, has no doctrines. It is a method and nothing more.

### *Philosophical Analysis and the Traditional Tasks of American Educational Philosophy*

I mentioned earlier that until the present decade there has been some reluctance on the part of educational philosophers to discover what worth conceptual analysis has for educational theory. They felt that the central tasks of the philosophy of education were to assess the meaning of the main currents and tendencies in the culture and to offer guidance on fundamental matters of educational practice on the basis of this interpretation. A cursory glance at the philosophical analysis movement revealed little of value for these central tasks.

Indeed, the older generation of educational philosophers not only seemed uninterested in using the new techniques of analysis themselves, but also seemed to feel hostile to their use by others. They felt, I suspect, that the prominence of analytical philosophy in departments of philosophy would pressure educational philosophers to redefine the aims of the philosophy of education in order to be able to use the insights of the analysts. And these older men, many of whom were deeply committed to the thesis that philosophy must be of some use in coming to understand ourselves and our culture and in improving our practice on the basis of this understanding, were unwilling to see the philosophy of education become as removed from the life of the culture as, in their opinion, academic philosophy had become.

Developments within the past decade have indicated that these worries were largely without foundation. As long as analytic philosophy remained a narrow academic enterprise it was able to sustain a certain measure of intellectual cuteness, a form of academic entertainment and self-deception. My colleagues in other disciplines will, I am sure, recognize this tendency closer to home.

The thesis that philosophy is nothing but the analysis of the expres-

sions of ordinary language does have great shock value, but unfortunately there is little else to be said in its favor. Indeed, there are some striking similarities between this thesis and the special philosophical theories which analytic philosophers so roundly condemn. Furthermore, the doctrine that philosophy has no doctrines is clearly self-contradictory. Thus, there is good reason for educational philosophers' rejection of much of the *program* of analytic philosophy. The rejection of the program does not, however, justify rejection of the new insights and techniques.

Analytic philosophers of education have shown little interest in philosophical programs. The original cuteness of the movement has not survived the voyage from Oxford and Cambridge drawingrooms to American educational research centers. Today one finds few educational philosophers arguing that philosophy is analysis, but many applying conceptual analysis in a wide variety of areas of theoretical and practical concern. It remains to be indicated how and where such analysis can contribute to these central concerns of traditional American educational philosophy.

One sort of interpretation of culture which has a clear and traditional claim to the title "philosophy" is the analysis of the concepts employed in the shared activities which constitute the life of the culture. To understand American culture one must understand the concrete forms taken by such activities as explanation and evaluation in addition to understanding practices of education, religious worship, and advertising. Part of this analysis would consist in an account of the concepts "to know," "to have a reason," "to have evidence," "to believe," etc. One may get some feel for this by considering what is to be learned about American culture in accounting for the various uses of the verb "to believe," distinguishing "believe that ....." and "believe in ....." contexts. In this sense, all conceptual analysis can be taken as cultural interpretation, as conceptual orientation to a set of shared activities as practiced in a particular time and place. (This idea is supported in Dewey, 1963; Winch, 1958; and Collingwood, 1940.)

Where the interpretation and assessment of culture extends beyond this analytical kind of orientation, it can make use of conceptual analysis. For example, in explaining the widespread disaffection of students in our colleges and universities, an explanation of the concepts of knowledge and rationality which are prevalent in these institutions would be a central part of any interpretation which had a valid claim to the label "philosophy" (Aiken, 1968). For it is this conceptual aspect of the culture which their special training and their knowledge of their history equip philosophers to discuss with special competence.

During the logical positivist reign of terror and virtue which dominated the middle period of the analytical revolution in philosophy, judgments of value were said to be meaningless nonsense. Some writers on education



have been too fearful to notice that most recent trends have been more gentle. The positivists had a theory about what these judgments *have to mean*; the recent conceptual analysts have carefully investigated this form of life and given an account of its conceptual structure. But perhaps it has not been sufficiently emphasized (at times it has been explicitly denied) that there is a practical point in coming to understand the conceptual schemes with which persons operate (emphasized in Murphy, 1965; denied in Moore, 1963). Philosophy, despite Wittgenstein's dictum to the contrary, does not leave everything the way that it is; or rather, it certainly should not do so if the professors know the good in what they are doing. The impulse to philosophy is complex, and often arises from practical difficulty, not only from context free "puzzlement" as some of the very cutest recent analytical philosophers have maintained. Sometimes men seek clearer understanding of their forms of life because those forms are in a mess and they want to stop and locate themselves.

Education is an enterprise which raises many problems that are of pressing importance and that are at least partially clarified through philosophical analysis. One need not maintain that the philosopher has a special and peculiar insight about the value of everything (and knows the value of horses better than a horseman) in order to recognize that he has something quite special to offer in understanding and dealing with values. The man with the map of the cave does not have any special senses, but he can help us to get around down there. Contrary to the underlaborer conception, the analytical philosopher of practical affairs can not, in general, limit his attention to abstract conceptual analysis. Under whom would this underlaborer labor? Few non-philosophers have shown either much interest in abstract questions of analysis or much willingness or ability to "apply" the insights of analytic philosophers. Let us grant that as intellectual labor is divided at present, it is not philosophers but substantive experts in areas of practical concern who have primary competence in making value judgments. If we believe that conceptual analysis can be of practical importance in clarifying substantive issues, then this hard and fast distinction between philosophers and substantive experts must be abandoned. We must be willing to see philosophers acquire relevant substantive expertise and experts acquire considerable philosophical sophistication. Analytical philosophers should continue to leave the "deep, unanswerable" questions at rest, but they ought to concern themselves with questions which will be answered in practice with or without their help, when they have something to add to the search for a secure and reasonable answer.

Philosophical analysis in isolation from other intellectual tools and the wisdom of the practitioner will solve very few interesting educational problems. Therefore, when educational philosophers allow themselves to become isolated from the rest of the world, they make the dictum that philosophy leaves everything the way it is true by their own conduct.



## *Specimen Philosophical Analyses*

This section contains some examples of problems clarified and resolved through philosophical analysis. The first example concerns the relation between teaching and learning.

### *Specimen 1: A Lesson About "Teaching"*

Many questions about the relation between teaching and learning are scientific questions and can be answered only by experiment. Even here philosophical analysis may be important. The meaning of an hypothesis, as well as the empirical data, affects the truth-value of the hypothesis. It will generally be scientific investigators, and not philosophers, who will operationalize the concepts in educational research. But this is a task which requires a sensitivity to the pre-analytic meanings and uses of expressions, and is one point in educational research which philosophers can focus on fruitfully. (See Komisar and MacMillan, 1968.) There are conceptual tangles here as well. Arguments arise about whether there can be teaching without learning, and the *can* is logical, not empirical. It is sometimes held that *A* has not been teaching unless someone has learned what *A* intended him to learn. Is this thesis correct? This question is philosophical because the solution to the problem it raises does not depend primarily on data. It is the interpretation of data that is in question. It is agreed that *A* attempted to get *B* to be able to do problems of kind *K*, and agreed that *B* can not do these problems, because he has paid no attention to *A*. Now, has *A* been teaching or not?

Those who answer in the affirmative argue that *in some sense or other* *A* must correctly be described as "teaching." Those who answer in the negative can point to the fact that it would be incorrect for *B* to say "*A* has taught me to do problems of kind *K*," for by saying that, *B* implies that he has learned how to do problems of that kind. Those speaking for the negative may add that while *A* has not taught *B* how to do problems of kind *K*, it is equally true that he has not taught him anything else either. The argument may conclude that a person simply can not be teaching . . . without teaching something. Hence, *A* has not been teaching.

The solution here, as so often in philosophy, lies in making a distinction between different uses of a word. Teaching, like law, is a recognized profession, and teachers engage in an activity called "teaching" which is largely defined by social context. Teachers go to schools, try to get students to learn a variety of things specified by the State and tradition, and are licensed by the State to engage in teaching. People are identified as teachers largely through these forms, and in one use of the verb *to teach*, anyone who is in this way identifiable as a teacher engages in the activity of teaching by virtue of that fact alone and without regard for the success

of his endeavors. It is in this use that "I teach" is an appropriate answer to the question, "What do you do for a living?" One might be tempted to call this the *professional* use of the verb, except that the professional teacher is only one among many who engage in teaching even in this sense (for example, Sunday School teachers, parents used as teachers in community schools, etc.). Because using the verb *to teach* in this sense carries no implication about learning, it can be marked as the *task* use of the verb. This can be contrasted with the *success* use, which does carry this implication. (I have largely followed Scheffler, 1960. The source of the task-achievement distinction is Ryle, 1949, pp. 149-153.) Of course, there would not be any teaching at all if there were never any learning *at all*, even if *teaching* is taken in the task sense only. That is, the application of the verb *to teach* in any familiar sense has the (empirical?) presupposition that someone at some time or another learned something. (The concept of teaching has been among the most discussed of educational concepts in recent years. See Komisar, 1968.)

Suffice it to say that when the verb *to teach* is used first in the task use and then in the success use, it is no contradiction to say that A has been teaching algebra for years, but has never taught anyone algebra. Clearly, this distinction has resolved the original problem, for with a clearer picture of the use of the verb *to teach* there is no longer any need to ask whether teaching really does imply learning or not.

### *Specimen 2: Some Means for Operating With "Aims"*

It is sometimes assumed that the aims of education are another special province of philosophy; that philosophers have some special method which enables them to know what people should aim at in their practical undertakings. It is as though the philosopher is both a hedgehog and a fox: he knows one big thing and many little things. He knows the will of God, or the nature of things, and from this he in some way deduces the aim of teaching seventh-year math. Alas, the situation is the same here as it was when the value of horses was under discussion. Philosophers simply are in no position to state helpful, concrete aims for seventh-year math. Yet some problems about aims arise because of failure to understand the conceptual structure of practical discourse, and here the techniques of analytic philosophy are useful.

To illustrate, I discuss one aspect of educational behaviorism. The widespread acceptance of this attitude lies behind much educational theory and practice. My remarks here are intended to show that one important argument type used by behaviorists must be modified or abandoned. It is through such modification that theoretical attitudes begin to lose their attractiveness and, hence, their practical significance.

Behaviorists frequently attack educational aims that are stated in non-

behavioral terms on the ground that these aims are vague, or grandiose. I argue here that at least sometimes the statements which are attacked in this way appear to be vague or grandiose only because they are forced into an incorrect pattern of interpretation due to a failure to understand the complexity of our aim-stating language.

Most statements of aims are naturally interpreted as specifying some terminal state or event to be brought about by appropriate but logically independent means. On this sort of interpretation, a determination as to whether or not the aim has been satisfied is made by reference to what takes place or is the case at the *end* of the teaching interval only. The word *aim* suggests objectives or goals of a process. Even if educational behaviorists and other educational researchers and practitioners allow that there are *process aims*, the natural inclination is to interpret statements of aims in the above manner.

Many mentalistic aims are vague when interpreted in the product manner. By saying that they are vague, I mean that empirically applicable conditions of aim satisfaction remain relatively undetermined by prior convention. A vague aim can be made clear by specifying an applicable criterion.

I now consider an example of a mentalistic aim which was been attacked, in the manner under examination, as vague. I argue that this aim, while vague on the product interpretation, is not vague when interpreted as a process aim. I then argue that there are significant practical differences between the original aim and the proposed behavioristic reductions. Suppose that a supervisor enters an algebra class and finds the teacher drilling the students on rote techniques for solving problems of a certain kind, without explaining why the techniques yield correct solutions or relating them to techniques for solving other sorts of problems. Some students plug away mechanically, while others become bored and restless. The supervisor suggests that the teacher should aim at getting the students to understand algebra, not merely to be able to do the problems on the final examination.

If this aim statement is interpreted as specifying some end state of the teaching interval, (which is, in this case, the whole algebra course) it would mean that at the *end* of the course, the students should not only be able to do the problems on the final examination, but in addition should have another property; they should understand algebra. The teacher finds it hard to get very much concrete guidance from this additional aim. He knows how to enable students to score well on the final examination, for he either controls the final examination or knows that the same kinds of problem appear annually on statewide examinations. But what would it be to get the students to understand algebra? As a terminal description of the students, that is very vague, and indeed, many



likely specifications of what this means in concrete terms make it a grandiose aim for ninth-year students to be expected to satisfy. Perhaps it means that they can figure out a technique for any algebra problem they can not solve directly, for example. At least, it is so vague that it is not at all clear that it does not mean this. The behaviorist is absolutely correct in arguing that this aim statement is vague and grandiose on this product interpretation (McNeil, 1964).

But if there is an alternative interpretation of the supervisor's aim statement, then the proposal to reduce this aim to behavioral terms is premature. One good reason for searching for an alternative interpretation is that *prima facie* it seems preposterous to deny that algebra teachers should aim at getting their students to understand algebra.

One alternative is to interpret the stated aim as applying to the process itself, and not as specifying a desired product. There may be some reluctance to accept process interpretations of aim statements because the common association of "aim" and "goal" makes one think that aims *have* to be products. But an examination of concrete cases should show that at least sometimes the process interpretation is more reasonable. Consider, for example, the aims of a maintenance man: to maintain cleanliness, efficiency, safety, etc. Interpreting the troublesome statement that the teacher should aim at getting his students to understand algebra in this way, it means that the teacher should see to it that throughout the teaching interval the students understand the concrete lessons, techniques, and principles which they are taught. Now, is this aim so vague or grandiose?

According to the above definition, a stated aim is vague to the extent that its satisfaction conditions are undetermined by prior convention. "Understanding algebra" as a product aim is very vague; but for the understanding of concrete lessons and techniques it is generally possible to specify non-innovative behavioral criteria. Behavioral criteria are test conditions which under normal conditions are indicative of aim satisfaction. Most teachers can tell whether their students understand the concrete lessons taught in schools. They do this by observing facial expressions and gestures in some cases, and also by contriving situations in which the students are unlikely to respond correctly unless they understand the lesson. The construction of these situations, in which understanding claims at this level (not the level of algebra-in-general) are tested and falsified, is one of the most common features of good pedagogy. Teachers utilize conventional behavioral criteria as a matter of course.

Conventional criteria are sometimes inadequate. Holt (1964) explained how appropriate facial expressions and gestures become unreliable indications of understanding school lessons. Because these are among the primary behavioral indications of understanding, poor teachers have been satisfied with these indications and have not tried to get beyond them,



have attempted to bring about these indications, and have reinforced them. The natural consequence has been that the children have learned to make these expressions and gestures in the absence of understanding, making school room conditions abnormal in the relevant respect; the gestures and expressions do not indicate understanding is there.

Educational researchers also rely upon conventional criteria in their specifications of more precise conditions. If they did not do so, the original aim statement, the object of respecification, "reduction," etc., logically could not provide guidelines for the behavioral or more precise new aim. Because educational theory and practice rely on prior convention in applying these aim statements, it follows that they are not vague in any clear sense. Once the criterion is specified, it will be seen that these aims are down to earth, not grandiose.

It may be argued that there is no practical difference between having behavioral aims and aims which have behavioral satisfaction criteria. If this objection were sound, it would mean that it would make no difference for educational practice whether mentalistic aims were salvaged from behavioristic criticism or not.

The answer to this objection is that were it possible to state exhaustive necessary and sufficient behavioral conditions for satisfaction of mentalistic goals, and not merely behavioral criteria with the normality qualification, then of course satisfaction of the conditions would logically entail satisfaction of the aim. If it is not possible to state exhaustive behavioral conditions, then there is a logical gap between bringing about the criterial behavior and satisfying the original "mentalistic" aim, a gap caused by the "normal conditions" qualification. This gap is of great practical significance, for aiming at ordinary criterial behavior will generally make conditions abnormal, in the respect that under those conditions the behavior is no longer indicative of original aim satisfaction. (Consider the point attributed to John Holt above.) Once criterial behavior becomes aim, short cuts skirting the original aim become available.

Because of limitations on the scope of this chapter, I will not argue, but merely assert that exhaustive behavioral conditions for most mentalistic aims can not be stated. (See Waismann, 1965; Hart, 1965.) Hence, in practice this gap remains.

### *Philosophy, Education and the Doomsday Threat*

I suggested earlier that as long as educational philosophers effectively isolate themselves from the rest of the world, they make the dictum that "philosophy leaves everything the way it is" true by their own conduct.

*"But what can be done? The world has gone mad, and the deep cleavages and unresolvable tensions in the world are not under the control*

of human agents, and least of all professors. This makes it pointless to try to change things for the better." With these words the prophet of doom reveals the dictum that philosophy leaves the world as it is, to be not mere entertainment but the very consolation of philosophy.

Well, what can be done? It should go without saying that no amount of conceptual analysis by itself is going to get us out of this mess. Yet some things are of unquestioned worth even here, and one of them is quality education which is geared to effective action in the world. Not the least important, nor the least traditional way in which the philosophic mind can contribute to the creation of quality education is to expose the thinking about education which it finds all around to searching philosophic criticism. (The classic work of this type is Bode, 1927.)

Philosophers of education must do what they promise; they must philosophize about *education* and the theories which their colleagues in the scientific and practical areas of educational research place before them. Questions that are really about education are rarely if ever straightforward problems of existing intellectual disciplines. For this reason I think the interdisciplinary concept used by professors in American Studies should be tried. Groups of students who are being trained in a deep and thorough way by the departments which represent the living embodiment of traditions of intellect should be meeting in seminars with professor of different and conflicting persuasions, and in these meetings all should be discussing questions of common concern about education. And in these seminars, the value of the insights of these traditions of intellect should be tested in a scientific (open, honest) way.

It is not modesty but philosophic confusion which makes professors of education assume that we in the educational arena must simply accept the products of reflection which are generated elsewhere in the university, and must be grateful for and uncritical of every morsel we are given. The value of these products for us is in large measure determined by what they have to contribute to deeper understanding of education, and these ideas come without credentials. I have suggested that some of the offensive aspects of analytical philosophy were removed by putting it to work. It is about time that we discovered what recent psychology, for example, really has to offer. And when we examine the products of recent psychological reflection, let us keep the following principles in mind. If these psychological theories fail to fit modest classroom facts, then we must reject the theories and not deny or ignore the facts. And if the psychological results are practically useful only to the extent that we falsify and distort our pre-theoretical conception of what education is and should be about, then it is not our conceptions of and ideals for education, but rather these theoretical reflections, which must generally be rejected. When the exigencies of practical endeavor do not set the context from theoretical

reflection on practice, such reflection is cut off from its ultimate condition of verification, and is theoretically irresponsible regardless of its scientific trappings and pretensions. (See Komisar, 1966; Peters, 1959, pp. 119-137.)

The comprehensive understanding of education will consist of more than a mere accretion of ill-fitting pieces from all over the intellectual map (except the school of education itself). This understanding can only be created if we all put what we bring to the study of education squarely on the line and do so together. It is only as we begin to do this that we can move from speculation to secure assessment of the role of philosophy in this task. Similarly, this speculation that the interdisciplinary seminar form will create effective communication remains to be tested and made more secure in practice.

I have argued that conceptual analysis and intellectual criticism based on this analysis, in a context of mutual respect and understanding among philosophers, educational theorists, and practitioners, will be the primary contribution which philosophers can make to the improvement of educational practice.

There is one more contribution which philosophers are specially equipped to make which may be of somewhat more immediate effect. They can try to understand the growing currents of unrest and uneasiness in the world, and interpret these to themselves, their colleagues, and in some cases, even to those who make these currents.

I began this essay by discussing some recent developments in educational philosophy. I have ended by trying to say something about the situation we all find ourselves in, and what it means to be a philosopher within this context. We live in a world which is torn by conflict and threatened with nuclear annihilation. Education offers one hope for moving into a future which is truly more secure and better than the present. This new education must be one which helps people to live with themselves and deal effectively with the world as well. It must be an education which fosters self-respect and which creates sympathies that run wide and deep. This is quite obviously not the best-selling brand. If in saying this I sound to some more like a philosopher than I have throughout, I can console myself with the fact that everyone must surely know that this is what education simply must become, and that no one needs philosophers to tell them this. But there is little that we philosophers can do to bring such education into being, because our distinctive contribution, heightened conceptual awareness, is of practical significance only when possessed by those who change the world in other ways, and by the time the shared attention is created which this requires, it may quite simply be too late.

## Bibliography

- Aiken, Henry David. *Rationalism, Education, and the Good Society. Studies in Philosophy and Education* 6: 249-81; 1968.
- Bode, Boyd H. *Modern Educational Theories*. New York: Macmillan, 1927.
- Camus, Albert. *The Myth of Sisyphus*. New York: Random House, 1961.
- Collingwood, R. G. *An Essay on Metaphysics*. Oxford: Oxford Univ. Press, 1940.
- Dewey, John. *Philosophy and Civilization*. New York: G. P. Putnam's Sons, Capricorn Books, 1963. Chapter 1, pp. 3-12.
- Flew, Antony (editor). *Logic and Language*. Garden City, N.Y.: Doubleday, 1965.
- Gorovitz, Samuel and Williams, Ron G. *Philosophical Analysis*. New York: Random House, 1966.
- Hart, H. L. A. The Ascription of Responsibilities and Rights. *Logic and Language*. (Edited by Antony Flew.) Garden City, N.Y.: Doubleday, Anchor Books A449, 1965.
- Holt, John. *How Children Fail*. New York: Dell, 1964.
- Huxley, Aldous. *Island*. New York: Harper and Row, 1962.
- James, William. *Pragmatism*. Cleveland: World, 1955.
- Komisar, B. Paul. The Non-Science of Learning. *School Review* 74: 249-65; 1966.
- Komisar, B. Paul. Teaching: Act and Enterprise. *Studies in Philosophy and Education* 6: 168-93; 1968.
- Komisar, B. Paul and MacMillan, C. J. B. (editors). *Psychological Concepts in Education*. Chicago: Rand McNally, 1968.
- MacMillan, C. J. B. and Nelson, Thomas W. *Concepts of Teaching: Philosophical Essays*. Chicago: Rand McNally, 1968.
- MacNeil, John D. The Influence of Programmed Learning on Curriculum Research. *Programs, Teachers, and Machines*. (Edited by Alfred de Grazia and David Sohn.) New York: Bantam Books, 1964. Pp. 185-92.
- Moore, G. E. Lecture. *Analysis* 1: 28-30; 1933.
- Murphy, Arthur E. *The Theory of Practical Reason*. La Salle, Ill.: Open Court, 1965.
- Passmore, John. *A Hundred Years of Philosophy*. London: G. Duckworth, 1957.
- Peters, Richard. *Authority, Responsibility, and Education*. London: Allen and Unwin, 1959.
- Peters, Richard. *Ethics and Education*. London: Allen and Unwin, 1966.
- Peters, Richard. *The Concept of Education*. London: Routledge and Paul, 1967.
- Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson's Univ. Library, 1949.
- Scheffler, Israel. *The Language of Education*. Springfield, Ill.: Charles C. Thomas, 1960.
- Scheffler, Israel. *Conditions of Knowledge*. Chicago: Scott, Foresman, 1965.
- Thomas, Sid B., Jr. Authority and Law in the United States, 1968. *Ethics* 79: 115-30; 1969.
- Urmson, J. O. *Philosophical Analysis*. Oxford: Oxford Univ. Press, 1956.
- Waismann, F. Verifiability. *Logic and Language*. (Edited by Antony Flew.) Garden City, N.Y.: Doubleday, Anchor Books A449, 1965. Pp. 122-51.
- Warnock, G. J. *English Philosophy since 1900*. London: Oxford Univ. Press, 1958.
- White, Morton Gabriel. *Towards Reunion in Philosophy*. Cambridge, Mass.: Harvard Univ. Press, 1956.
- Wilson, John; Williams, Norman; and Sugerman, Barry. *Introduction to Moral Education*. Baltimore, Md.: Penguin Books, 1967.
- Winch, Peter. *The Idea of a Social Science*. London: Routledge and Paul, 1958.
- Wittgenstein, Ludwig. *Philosophical Investigations*. New York: Macmillan, 1953.

## AUTHOR

LEONARD JOSEPH WAKS Address: Stanford University Title: Assistant Professor of Education and Philosophy Age: 26 Degrees: B.A. and Ph.D., Univ. of Wis. Specialization: Philosophy of Education.



# BEHAVIORAL SCIENCE

## An Interdisciplinary Journal

**BEHAVIORAL SCIENCE** publishes original articles which develop theories of human behavior to explicate basic behavioral concepts and report data specifically oriented toward such theories. The editors are especially interested in the submission of articles with a broad interdisciplinary approach which would not normally be found in journals devoted to a single discipline. Applications to the mental health field, within this framework, will receive special attention. The journal is bimonthly, appearing in January, March, May, July, September, and November.

### SELECTED CONTENTS OF RECENT ISSUES

- "Behavioral Modification of an Autistic Child" by *Eleanor R. Brawley, Florence R. Harris K. Eileen Allen, Robert S. Fleming, and Robert F. Peterson* (Vol. 14 [2], March, 1969)
- "Schizophrenia: Diagnosis and Prognosis" by *Gabriel Langfeldt* (Vol. 14 [3], May, 1969)
- "Toward the Study of Biopolitics: A Cross-Sectional Analysis of Mortality Rates" by *Michael Haas* (Vol. 14 [4], July, 1969)
- "Primate Frontal Lesions and the Temporal Structure of Behavior" by *Waller Tubbs* (Vol. 14 [5], Sept., 1969)
- "Equilibrium, Optimum, and Prejudices in Capital Markets" by *Karl Borch* (Vol. 14 [6], Nov., 1969)

### COMING IN VOLUME 15, 1970

- "The Construction and Calibration of Behavioral Rating Scales" by *Bruce A. Barron, Jules Hirsch, and Myron Gluckman*
- "The Irrationality of Transitivity in Social Choice" by *Peter C. Fishburn*
- "A Consideration of the Neural Basis of Intelligent Behavior: Comparative Studies" by *P. J. Livesey*

and

The first in a series of annual symposium issues:

#### **Systems Principles in Biology**

with Eugene Ackerman, Britton Chance, Frederick A. Dodge, Jr.,  
Fred S. Grodins, Barbara E. Wright, and others

**BEHAVIORAL SCIENCE**  
**Mental Health Research Institute**  
**The University of Michigan**  
**Ann Arbor, Michigan 48104**

*Subscription rates: Individuals, \$15.00; libraries, \$30.00; single issues, \$3.00*

Make checks payable to **BEHAVIORAL SCIENCE**

## 6: POLITICS OF EDUCATION

MICHAEL W. KIRST\*

*Stanford University*

EDITH K. MOSHER

*University of Virginia*

The politics of education is a new and still largely uncharted area of research concentration. As recently as 1959 Thomas H. Eliot, a professor of political science, presented the need for a beginning of research in this field. Since then a significant number of research studies have been published and new courses instituted at universities. Researchers recruited from political science and education are debating the proper scope, methods, and objects of inquiry. In 1969 AERA had enough interested members to form a special interest group to further such efforts.

These developments emerge from a growing realization that educational problems and issues call for "policy-science" modes of analysis.\*\* Since World War II, many social scientists have accepted the policy-making processes as intrinsically suitable objects of study, and it has gradually become apparent that political issues were either neglected or inadequately interpreted by those scholars known as the philosophers, historians, sociologists and economists "of education."

What is distinctive about a policy-science orientation in educational research? Studies of this genre have been concerned with three issues: 1) the allocative effects of political actions—who gets what, when, and how; 2) the nature of the present educational-political system and changes in its characteristics over time; these concerns of political systems analysis are discussed below in Section four; 3) the effectiveness of the system in reaching what researchers have deemed to be desirable policy objectives—e.g. innovation, increased fiscal support, community participation, etc. Most of the research has treated questions of allocation: How are decisions made?

\*Drs. Alan K. Campbell, Syracuse University, and James Guthrie, University of California, served as consultants to Drs. Kirst and Mosher on the preparation of this chapter. The authors gratefully acknowledge the editorial help provided by Dr. George Wren, University of Virginia.

\*\*The term came into common usage with the publication of Lasswell and Lerner's *The Policy Sciences* (1951). Scholars prominently associated with this orientation to study of public policy and political behavior include Paul Appleby, Edward Banfield, Robert Dahl, John Gaus, Bertram Gross, V. O. Key, Paul Lazarsfeld, Charles E. Lindblom, Robert Lynd, and David Truman.

Who exerts influence upon decisions about budgets, curricula, facilities, desegregation, etc.; and upon whom are those influences exerted? Who shares in the benefits of political activities? Under what conditions is community control of schools and professional educators maximized?

As an emergent area of study bridging divergent research traditions in political science and education, the politics of education is affected by the characteristics and limitations of both environments, past and present. The methodology of the politics of education is as yet unstable and untested. It is important to bear in mind that this is a fledgling area of research that has not yet developed a distinctive empirical base and methodological sophistication. There are at present, however, clear signs that this will occur.

### *Patterns of Established Thought: The Legacy of Closed-System Politics*

Neither of the parent disciplines—political science or education—has provided clear concepts nor ready-made, tried-and-tested methods for the study of educational politics and policy formulation. Political science suffers from a lack of agreement on methods and from the absence of widely accepted theoretical frameworks. Truman (1965, p. 869) spoke of the 1960's as the era of the "quest for a discipline" within political science:

Given the looseness and especially the lack of precision in the prevailing implicit agreement on what to do and how to proceed in the field, its weakening and gradual dissolution were bound to be followed by a confusion of competing and divergent, if not incompatible, views of the appropriate questions to be asked and the proper methods to be used. How long that state of affairs is likely to exist is anyone's guess.

Truman (1965, p. 866) described the professional consensus among political scientists that dissolved after World War II as including: 1) an unconcern with the properties and requirements of political systems, 2) a neglect of theory, 3) a consequent enthusiasm for "science" that rarely went beyond raw empiricism, and 4) a confining commitment to concrete descriptions.

In the early 1900's, the reform of school administration that sought to divorce education from overt political conflict affected also the nature of research activities. Prior to the late 1950's, the governance of the U.S. educational system received relatively restricted and low-level scholarly attention. Vastly outnumbered by their colleagues in the psychological subfields, researchers pursuing such interests were almost exclusively in schools of education, affiliated professional associations, state departments

of education, or the USOE. Their studies had methodological limitations comparable to those which Truman attributed to traditional political science research.

Common to all governmental research are certain basic difficulties in developing well-defined concepts and fruitful hypotheses to guide the collection and ordering of complex masses of data. In education a professional parochialism that has been widely documented was a further deterrent to scientific results. For example, Martin (1962), a political scientist who studied educational politics in the early 1960's, asserted that the suburban school system was embedded in a universal and tradition-laden system which over the course of a century achieved liturgical status. In essence the liturgy was: education is a unique function of government that must have its own separate and politically independent structure. The administration of education should be uninvolved in "politics" and professional unity should be the norm.

Apparently, what Sayre (1963) termed "the schoolmen's political myopia" and Iannaccone (1967) called "the politics preferred by pedagogues," guided the selection of research problems. This presumed relationship between professional ideology and a dearth of inquiry concerning the political aspects of education warrants close examination. Iannaccone stressed that school administrators were highly successful in attaining their goal of "boundary maintenance," i.e., in freeing themselves from external political constraints to a greater degree than did the directors of most other public programs. Even though school systems are highly diverse in size and wealth, and represent what Sayre called the "last stand of localism" in American government, they have a high degree of uniformity and stability in their basic organization, functions, and management policies. When thus characterized by political isolation, tight boundaries, and a slow rate of internal change they exemplify social systems generally defined as "closed systems." Iannaccone concluded that for most of the twentieth century the politics of education typified "closed-system" politics.

He did not mean, however, that "education is not (or was not) politics," as educationists long claimed. However, there was no incentive for scientific research to question basic educationist tenets and appraise existing power structures. Like the officially sponsored researchers concerned with issues in administration, instruction, finance, and curriculum, students of the government of education paid their way by seeking answers to questions of urgent interest or importance to school administrators. Most studies dealt largely with internal and stable aspects of educational institutions and practices; they were delimited in scope to specific program areas, educational levels, states or localities, and so on. A considerable crop of books, monographs, and dissertations concerning school law and finance, administrative regulations and practices, and judicial decisions treated these



subjects descriptively, often in exhaustive detail and comparative classifications, but seldom analytically or critically. In studies of community teacher, and student populations researchers used pre-World War II non-analytical models of sociologists. Other less empirically oriented researchers continued to employ either historical or "arm-chair" modes of inquiry and to state abstract, global and sometimes prescriptive conclusions concerning the goals and properties of educational enterprises. For a brief but perceptive summary to the "State of the Art" in past decades see Ralph W. Tyler (1965). In none of this research production does one find much attention paid to the political behavior of those who ran the schools, to the relation of schools to other municipal governmental agencies, or to involvement of citizens in school policy-making.

The value that educators attached to the isolation of their activities from politics may, to some degree, explain the absence of research comparing public school systems with other social institutions having education-related goals. It was an economist (Machlup, 1962) who treated public schools as only one segment of the country's resources and institutions for knowledge production and transmission. However, educators showed considerable concern with developing stable and uniform measures of *social inputs* to educational systems—money, personnel, and students. Educators almost totally neglected the related task of developing comprehensive and similarly concrete measures of the quality of the *system outputs* of schooling. As late as 1964, those conducting an extensive NEA-USOE study of school "drop-outs" were still working on the problem of standardizing the definition of the term (Schreiber, 1964).

The tacit assumption shared by administrative researchers and their professional colleagues was that levels of educational quality and school performance would vary directly, positively, and almost exclusively with levels of financial support provided by the taxpayers. They attempted to develop school aid formulae and rationales that would bring in more money for schools. The notable school finance surveys and research reports of Mort and his colleagues were thus strongly conditioned by an unacknowledged political goal—larger budgets (see Benson, 1968). One of the most extensive research projects of the 1940's exemplifies this bias. It distilled seventy pieces of research concerning the adaptation process in education and the adaptability of school systems. Its conclusions are closely tied to establishing the desirability of local district autonomy and increased fiscal support. Much attention was given to measures for building community support for the schools. (See Ross, 1951.)

The overriding importance of fiscal concerns is evident also in many of the most competently done and comprehensive of recent studies. When political scientists developed an interest in state and national educational politics in the late 1950's, they focused on the effectiveness of educators

in using political techniques and in penetrating governmental systems to obtain increases in state and Federal school aid. (See Bailey et al., 1962; Masters et al., 1964; and Munger and Fenno, 1962.)

### *Forces for Change—the Mid-Sixties*

It is almost a social science axiom that institutions or organizations that function as closed systems have great difficulty in carrying out the processes of objective self-examination for reform and self-renewal. The impetus for adaptation to changing contexts and demands almost invariably originates externally, and this appears to have been what happened in the beginning of research in the politics of education a decade or so ago. There is no need to catalogue here the social developments that have overwhelmed schools and professional educators. A number of environmental influences, however, are of particular importance to understanding the breakthrough in research preoccupations.

At the working level, a variety of social scientists from other disciplines, especially from political science, provided a significant stimulus to research in the policies of education. Funded in most cases by USOE or private foundations, they brought their heterogeneous modes of inquiry to the study of educational issues and enterprises. University centers for interdisciplinary and collaborative research were established at Syracuse, Oregon, Claremont, Harvard, Stanford, Washington University in St. Louis, City College of New York, and the Midwest Administration Center at Chicago. A growing number of papers appeared in diverse scholarly journals and, for the first time, several professional associations scheduled sessions on the politics of education.\* Collections of seminal research reports, with interpretive commentaries, are now becoming widely available in book form. (See Bendiner, 1969; Campbell et al., 1969; Gittell and Hevesi, 1969; Rosenthal, 1969a; Kirst, in press.)

The surge of academic interest in the study of educational policy can be traced, in some degree, to the changed attitudes of politicians. The prodigious post World War II demand for all kinds of domestic public service resulted in intense competition for fiscal support among governmental agencies and interest groups. Caught in the crossfire, legislators were no longer willing to accept the justifications educators continued to offer for their increasingly open-ended cost estimates of their needs. Legis-

\*Among such journals in disciplines other than education are: *Administrative Science Quarterly*, *American Political Science Review*, *Public Administration Review*, and *The Public Interest*. In education, several recently established journals frequently publish relevant research: *Educational Administrative Quarterly*, *Education and Urban Society*, and *Urban Education*. Journals devoted exclusively to educational research have begun to alert their readers to the importance of the new sub-specialty. See, for example, Schultz (1969) and the Summer 1969 issue of *Journal of Research and Development in Education*, entitled "The Public Interest."

lators began to ask for more meaningful evaluations of the results of formal schooling and to compare the potential contributions of education, broadly defined, to such competing policy and fiscal alternatives as manpower training, community action, food, and negative income tax provision.\*

Far-reaching proposals to change the traditional relationships among the three levels of educational government written by such eminent scholars and policy advisors as Roald Campbell, James Conant, John Gardner, Walter Heller, and Francis Keppel were widely circulated and discussed across the country. At the local level, some of the public concern with educational policy took on a new focus—that of redistributing political influence previously held by the professionals. Parents, teachers, minority groups, and students, previously powerless, went into action. In various locales they sought to obtain places on school governing bodies, mandates for consultation with officials concerning school policy, and decentralization of decision-making to the level of the individual school.

The Federal school aid legislation of the 1960's took some account of these trends by creating a number of watchdog citizen advisory panels and by calling for consultation by school authorities with community groups.\*\* The research sponsored by such advisory groups on the problems of curriculum change, school desegregation, educational opportunity for disadvantaged groups, and community involvement in school has shed considerable light not only on school practices, but also on political structures and processes that were previously known only to top level education officials and advisors. These intensified concerns of legislators and various publics in the schools had at least two immediate and obvious effects: 1) relevant data became both more abundant and accessible and 2) researchers attacked problems not previously studied, such as political influence and the representative character of school boards.

#### *IV: Conceptual Frameworks and Methodologies for Political Analysis—the Late-Sixties*

##### *A: The Impact of "Open-System" Politics*

The evidence that earlier forms of "closed system politics" restricted research supports the inference that the new open political context would be a far more unstructured and complex field of scientific inquiry. This

\*In 1965, Senator Robert Kennedy evinced so much concern about the potential productivity of federally-aided school programs that the unique evaluation requirement which was included in Title I of the Elementary and Secondary Education Act of 1965 was known informally as the "Kennedy Mandate." For a comprehensive and exceptionally interesting review of legislative policy-making for domestic programs during the 1950's and 1960's, see Sundquist (1968).

\*\*The U.S. Commission on Civil Rights and the National Advisory Council on the Education of Disadvantaged Children are good examples of citizens oversight groups.



has proved to be true and the current groping of researchers for appropriate strategies echoes the uncertainties which the public, the politicians, and the educators face in moving into "open-system politics." Lindblom (1968, p. 3) described public policy-making as marked by "complexity and apparent disorder"; his statement is equally applicable to the present study of educational politics.

Highly eclectic sets of research problems, conceptual frameworks and research techniques abound. However, one coherent and distinctive influence on research activity is discernible—the application of systems analysis models. In this context, "politics" is defined by one of the leading systems theorists, David Easton (1965a, p. 50), as the "authoritative allocation of values which have become widely desired and the explicit goals of public policy." In Easton's treatment of the structure of political systems, the term "authoritative" refers to the governmental decisions, policies, and activities which are the "outputs" of a system. "Values" and "goals" refer to the social demands and resources which, as *inputs*, create and maintain the system. (See Easton, 1965b.) The term *allocation* needs a detailed explanation in each research undertaking because it refers to the intricate interplay of particular ideologies, individuals, institutions, and interest groups. Since it is postulated that demands for authoritative action are divergent and shifting and that resources are generally inadequate to satisfy all the demands, the study of allocation processes in a political system is also the study of competition and settlement, or the avoidance of conflict.

The researchers who share the concern of Easton and others with analyzing the allocative process have focused on diverse and sometimes overlapping aspects or categories of contemporary political phenomena (e.g., legislative processes, administrative policy-making, community dynamics, institution-building educational innovation, school desegregation, school board and bond elections, and teacher militancy). Their investigations still tend to study in depth a single level of government, with a restricted set of variables, or with a small number of units. (See Gittell and Hollander, 1968; Bailey et al., 1962; Masters et al., 1964; James et al., 1963.) The studies highlight different components of the educational-political system, but the units of analysis are not yet numerous or broadly enough defined to be fully representative nor are the research designs sufficiently similar to provide for cumulative or comparative findings. Even within greatly restricted dimensions, there have been few studies dealing with the functioning of political processes over time or with the relationship between the various levels and branches of government. Colton (1969) pointed out, for example, the lack of theory and meager data that can be used to identify or explain the conditions under which local schools respond to state directives. Only recently has a major research study been completed in another long neglected area—the relationship between elementary-secondary and higher education at the state level (Usdan et al., 1969).



Alan Campbell (1968) emphasized that political scientists overlook the role of the judiciary in their analyses of educational policy-making, yet its significance is self-evident. There are several recent works which describe the legal-constitutional aspects of public school desegregation, but a dearth of those which analyze the impact of court decisions on the other political systems for governing education. Most investigators have followed the traditional practice of treating the operations of judicial systems as "given" or "constants" rather than as variables which have differential effects on other governmental systems. A notable exception is a recent study by Orfield (1969).

In short, little agreement exists about priorities or theory to guide research. This is a common deficiency of a new field of inquiry, but in this instance, it should not be attributed to the insensitivity of researchers to the requirements for scientific productivity. The limitations arise from the need to delineate and analyze vastly complex data, from the scarcity of funding and competent researchers, and from the use of diverse conceptual mappings of the field. As was the case in the early discovery and colonizing of North America, a variety of explorers have staked out and laid claim to sections of the unsettled terrain.

*B: The Influence of Imported Theories of Political Behavior on Educational Research*

The researchers who broke away from the earlier modes of studying educational governance have not followed the orderly progression of stages commonly attributed to scientific activity. That is, they have not moved from developing a hypothesis about some set of events to the task of empirically testing the hypothesis. Instead they have selected "targets of opportunity" for research based on their perception of pressing educational problems, available funds and talent, and access to data. Little attention has been given to the development of original conceptual frameworks; instead concepts and constructs have been imported from various social sciences, and these have tended to suggest rather than control the research design. A rather prevalent tendency to formulate objectives and conduct investigations in step-by-step sequences approximates what Glaser and Strauss (1967) envisaged as requisites for the development of "grounded theory." Masters et al. (1964, pp. 10, 11) stated the scientific aspirations of three researchers in launching comparative case studies of three states:

It is hoped that if and when our hypotheses are integrated with findings from other studies, they will take us a little further down that rough-hewn path that leads to a general theory of politics, or that stage in scientific development where research is conducted with a conceptual scheme or model that, for research purposes is widely accepted within the community of scholars.

Easton (1965b, p. 474) stated that the present body of research findings about political behavior should be regarded as providing the basis only for "partial theories of allocation." Schoettle (1968) analyzed a number of policy studies and listed "leverage points" for this type of research. Her schema might appropriately serve to categorize most of the relevant political studies in education. In general, they reflect what Schoettle described as a growing consensus among political scientists: namely, that purely legal and institutional descriptions can not adequately explain policy-making processes, and that no single focus is sufficiently comprehensive to provide a model of the process. Studies in the politics of education which illustrate the various categories are cited below. It should be understood that many other studies would have been equally appropriate for the purpose.

1. *Citizen perceptions, opinions, attitudes and voting practices with regard to educational issues.*

*Example:* The studies by Carter (1960) and by Carter and Sutthoff (1960) of voter attitudes toward their schools and of school practices that might lead to greater understanding among community electorates. Under this category studies, such as Greenstein (1965), of social and educational influences on political attitudes might also be included.

2. *The habitual strategies, information sources, and perceived self-interest of relevant policy-makers.*

*Example:* The study by Pois (1964) of the type of personal, political, and professional dominance which characterized Benjamin Willis's terms as Superintendent of the Chicago school system.

3. *Roles in the policy-making process; that is, the interrelatedness and interdependence of participants in various systems of action.*

*Example:* The study by Gross et al. (1959) of the role orientations of school superintendents and school board members.

4. *Premises and context for authoritative decision-making.*

*Example:* The study by Crain and Street (1966) of school integration issues in eight large city school systems of the North.

5. *Groups as interacting pluralities of individuals who share common political interests and goals.*

*Example:* The study by Rosenthal (1969b) of the political characteristics of contemporary teacher groups.

6. *The structure and impact of governmental institutions and processes.*

*Example:* The study by James et al. (1963) of the budget processes in fourteen large city school districts.

7. *The influence of elites or community power structures on educational policies and programs.*

*Example:* Kimbrough's (1964) study of informal power structures which influence educational policy in four southern counties.

C. *The Study of Politics and School Administration*

Earlier studies of school administration that ignored the influence of political variables were predicated on assumptions widely shared at the time by organizational theorists and students of public administration. That is, policy-making and administration were generally considered to be activities which were properly performed in separate contexts, by different officials. Researchers sought to develop comprehensive and scientifically tested principles of administration which would have application in all types of organizations. A classic example is Gulick and Urwick (1937). During the post World War II years, this objective was challenged by researcher of policy-science orientation who argued persuasively that the administrative functions of public officials were interpenetrative. Such views had a great influence on subsequent research in public administration. (See Appleby, 1949 and Peabody and Rourke, 1965.)

In education, however, the effort to study the effects of school bureaucracies on policy formulation and implementation lagged, perhaps because the influence of sociological theories or organizational behavior dominated much of the recent research in educational administration. (See Bidwell, 1965.) In the studies of local school district politics, the tendency was to consider the superintendent and the school bureaucracy as an entity, and to compare its behavior with that of the school board, the community, or the teachers' organizations. A frequent finding is that the "administrator" has decisive influence on most important matters of school policy most of the time. This research focus obscures the political roles of the associate and district superintendents, curriculum supervisors, and department heads. In many areas, particularly in curriculum policy, the reputed power of the superintendent may actually be exercised by his subordinates, so that he is himself unable to bring about changes in the system (see Gittell et al., 1968).

One promising strategy for inquiry concerning the overlapping concepts of "organizational-political" behavior was used by Rogers (1969) in recent research on the effort to desegregate the New York City schools. He studied the process of implementation from the school board level down through the successive layers of the hierarchy to the level of the classroom teacher. He employed extensive and ingenious methods to document the changes made by personnel in the central office subdivisions and in the field units. The Bailey and Mosher (1968) study of the first year of implementation of the Elementary and Secondary Education Act of 1965

was also concerned with politico-administrative events, principally at the national level. They traced the formulation of certain ambiguous congressional policy decisions and the subsequent efforts of federal administrators to interpret and put them into effect.

#### D: *Methodologies*

A more free-wheeling, wide-ranging attack on the understanding of the politics of education brought researchers face-to-face with the need to acquire greater methodological sophistication. They began to confront basic questions of scientific inquiry: What are the "telling" questions that will facilitate analysis of and generalizations about complex phenomena? What time dimensions must be explored? Should one select a few variables and study them in large populations, or consider the pattern of interacting variables in a single unit or system? In designing his own course of action no researcher can avoid choosing, deliberately or unconsciously, among a great many alternative procedures. Whether he acknowledges it or not, the initial choices channel his subsequent research. For example, in much of contemporary social science research, the difficulties of sorting out the effects of antecedent conditions on subsequent events, combined with the admiration of many researchers for the precise methods of the physical scientists, has reduced the number of studies of developmental and change processes. This concentration on historical studies characterizes the field of educational politics at present, though some studies might in the future provide the basis for badly needed longitudinal research. An almost unique example of a longitudinal study is that of school politics in Jackson County by Goldhammer and Farnier (1964) and Goldhammer and Pellegrin (1968).

Researchers in the field have tended to polarize around two basic approaches for sampling and selection of variables, both of which have advantages and disadvantages.

Some researchers favor a *survey approach*, in which a restricted set of variables, generally those susceptible to numerical measurement, are isolated and accepted as indicators of more general concepts. The variables are then studied in a relatively large and representative sample population. For example, "average annual per pupil expenditure" is a common measure of educational quality in a school district, and "voter turnout in local elections" is a measure of the degree of community concern with public services. The degree of relationship between these, and usually other similar numerical indices, is computed in a large number of school districts of varied characteristics.

Other researchers favor a *case approach*, in which all relevant educational-political characteristics of a community, institution, interest group, or set of events, are studied in depth. A case study may be as narrow



in scope as a single school board session or as comprehensive as the school system of New York City. The number of variables and complexity of relationships to be treated in a case of even modest scope becomes great. Much use is made of non-quantitative data derived from interviews with informants and researcher observations.

### 1. *The Survey Approach*

Survey methods permit the use of descriptive and inferential statistics, including, for example, the application of cluster analysis to the study of legislative roll-calls on education bills. Large samples theoretically reduce errors. Surveys are more easily replicated by other researchers, and data collection and analysis workloads may often be readily delegated to research assistants. However, there are risks that the basic assumptions of the statistical models are not satisfied, or variables used as indicators are inappropriate or insensitive. Investigators such as Campbell and Sacks (1967) have used multiple regression analysis to determine why some localities spend more money for schools than others and have a higher tax effort. Almost without exception the investigators found that the most important determinants of expenditures for education at local levels are economic variables. When Dye (1967) included political variables in the regression analysis, they were not significant. Those knowledgeable about local politics suspect that this technique merely obscures the real effects of political processes in determining tax and expenditure levels, and that additional "in-depth" forms of research are required (Zeigler and Johnson, 1969). Sharkansky and Hofferbert (1969) used factor analysis of political variables instead of multiple correlation and found that where a state is wealthy and shows high voter turnout and intense interparty competition, it is likely to score high on the level of welfare and educational services.

Extensive surveys established that state school aid formulae have long discriminated against urban school districts. It might appear that the recent reapportionments of state legislatures mandated by the U.S. Supreme Court would increase the number of representatives from urban areas and lead to some correction of the imbalances. But, in fact, suburban school districts have increased their relative proportion of state legislative membership more than have the cities, and the failure to act on the urgent priority of urban educational problems and needs persists on the state level. No research has been undertaken on the political processes underlying this situation (Campbell, 1966).

### 2. *The Case Approach*

Practically all phases of a case study require a high order of research competence, since this method is plagued by the risks of low reliability and sampling error. However, it is often argued that findings about the

patterning of variables, based on even a single case, hold more validity than findings concerning those same variables when they are treated in isolation from patterned interactions. James Fesler, a noted political scientist, concluded that the case studies issued by the inter-University Case Program are very valuable because they are not limited to mere chronicling (see Bock, 1962). Instead they provide an analytic understanding of the decision process, the organizational and political framework, and the substantive policy problems to be found in a "slice of government life." For somewhat similar objectives, the case approach was frequently used in the pioneer studies of educational politics. Since the studies were, in general, undertaken without any prior agreement on a common conceptual framework, the generalizations which they suggest must be considered to be tentative and partial. (See Masters et al., 1964, pp. 10, 11.) Given the embryonic state of political research in education, the existence of 50 states and 17,000 local school districts, the influence network of general governments, professional associations, and citizen groups, it is not surprising that researchers tended at the outset to concentrate on delineating the particular and the unique, rather than discovering "general principles."

Investigators are increasingly trying to combine the advantages and minimize the limitations of the two approaches, either by first conducting a series of pilot case studies as a basis for selecting variables which may, with greater confidence, be treated by survey techniques, or by using surveys of prior studies to select a set of variables about which to structure several in-depth case studies (James, 1966, pp. 40-42). Both of these approaches are a promising beginning but only a beginning.

For example, Rosenthal (1969b) used prior studies and theories to select four categories of independent variables (organizational strength, organizational opportunities, organizational behavior, and organizational influence) and four dependent variables as "policy outputs" (salary, personnel, curriculum, and school organization) for his study of nine teacher organizations in five cities.\* However, he relied in large part on the "reputational technique" whereby leaders of teachers' organizations attributed political influence to various participants in the policy-making process. This technique, used by Floyd Hunter and others in the identification of community power structures, has not been largely discredited on the ground that such elites could not be identified solely by associational roles or socio-economic status (see Schoettle, 1968).

### E: *Systems Analysis and the Study of Politics*

Since there is an increasing use of systems analysis in social science, especially in political science, it is not surprising that its potential value

\*Robert Crain (1969) explains a somewhat similar use of typologies by a research team which studied the politics of school desegregation in eight Northern cities.

should also be under examination by the students of educational politics. In a paper provocatively entitled "Political Systems Analysis in Educational Administration: Can the Emperor Be Clothed?" Sroufe (1969) asserted that the systems approach has introduced several studies—provided, in fact, a sort of window-dressing for them—but that it has yet to prove rewarding as a *method* of study. He did concede that the systems model could be used to "suggest the larger canvas, the total picture of which one's study is but a piece." Also, it provides a way of ordering complex situations to make them manageable for analysis, and it generates questions about the actors in the system. His position is consistent with that of Kaplan, who wrote in another context (see Almond and Coleman, 1960, p. 30):

Perhaps the first thing to be said about systems theory is that it is not a theory. It consists of a set of concepts. No propositions about the real world can be derived from infinitesimal calculus, or from the methods of science in general. Advice to a political scientist to use systems theory to solve a problem, even when it is the appropriate methodology, would advance him as far but no farther than would advice to a physicist to use the methods of science.

Such controversy aside, it is for the much-needed tasks of "mapping the field" that systems theory may best serve to advance further research in the politics of education. Iannaccone (1967), as well as Bowles (1968) and Decker (1969), paved the way by using its terminology and concepts to "paint a large canvas" of theoretical propositions. For example, Iannaccone stated that the government of education is characterized by long periods of stability interspersed with shorter periods of *abrupt* change. He attributed this phasing to the closed-system form of politics that does not generate alternatives to the policies advanced by educators. Since there is no "two-party system" in operation to create and institutionalize opposition to the *status quo*, changes in policy precipitated by external crises tend to be more abrupt than in fields characterized by open-system politics. (See also Campbell, 1968, who discussed the "periodic crisis" characteristics of educational government.) It should be possible either to prove or disprove these statements by means of appropriate empirical studies.

Iannaccone also discussed the relationships between organized interest groups and the educational policy "outputs" of state legislatures; he suggested that there are four basic structures of state educational politics. He then analyzed the findings of several research studies to see if and why states may change from one of these types to another over a period of time. He candidly admitted that the problems of generalizing from case study data about only eleven states, and from studies that differed in design, made his analysis highly speculative and tentative. However, one of his aims is to highlight the need for research of a more systematic type; and



he was able, by reference to a series of related studies of local school districts, to demonstrate their potential power for developing broad generalizations.

### *The Past and a View of the Future*

The emergence of the politics of education as a field of inquiry is associated more with developments in educational policy than with distinctive research techniques, methods, or theories. In the past research was retarded by practicing educators who posited that public education should be a unique nonpolitical function with a separate and independent governmental structure. This isolationist stance led researchers studying school operations to ignore their political environments.

In the past decade social forces affecting the schools have changed this situation and created a more open system of educational politics. Researchers now examine both the interchanges between the government of education and its environment, and the intra-system networks for influence. Most of the research has concentrated on the allocative consequences of political interactions—who get what, when and how.

Although this recent research is a promising beginning, we must move from micro-analysis to macro-analysis. There are 17,000 local districts and fifty states, and the government of education in each local district is to a degree unique. In many respects, these complex and differentiated entities do not lend themselves to broad scale statistical surveys. However, case studies do not permit us to generalize. We need appropriate combinations of both methodological approaches, and collaborative projects where several political researchers in different sections of the country use a common agenda, a common system of reporting, and a single research plan. Such a scheme might provide us with reliable findings simultaneously in all sections of the nation. Our efforts should, however, not be limited to a cross-sectional snapshot at one point in time. We need systematic study of change in educational government.

It would be helpful if theories could be borrowed from political science to guide our efforts. As Landau (1968) pointed out, however, this discipline is marked by a "high information level and low theoretic yield." For the most part, theory in political science has come by way of transfers from a wide variety of sciences in the hope of systematizing the field. Consequently, although our goal should be macro-analysis, many of the needed tools for the job are not at hand.

### *Bibliography*

- Almond, Gabriel and Coleman, James (editors). *The Politics of Developing Areas*. Princeton, N.J.: Princeton Univ. Press, 1960.  
 Appleby, Paul. *Policy and Administration*. University: Univ. of Ala. Press, 1949.  
 Bailey, Stephen K. and Mosher, Edith K. *ESEA: The Office of Education Administrators a Law*. Syracuse, N.Y.: Syracuse Univ. Press, 1962.



- Bailey, Stephen K. et al. *Schoolmen and Politics*. Syracuse, N.Y.: Syracuse Univ. Press, 1962.
- Bendiner, Robert. *The Politics of Schools*. New York: Harper and Row, 1969.
- Benson, Charles S. *The Economics of Public Education*. Second edition. Boston: Houghton Mifflin, 1968.
- Bidwell, Charles E. The School as a Formal Organization. *Handbook of Organization* (Edited by James G. March.) Chicago: Rand McNally, 1965. Pp. 972-1022.
- Bock, Edwin A. (editor). *Essays on the Case Method in Public Administration*. Brussels: International Institute of Administrative Sciences, 1962.
- Bowles, Dean. *Policy: Theoretical Research Implications*. Paper presented to the American Educational Research Assoc., Feb. 1968.
- Campbell, Alan K. *The Socio-Economic, Political, and Fiscal Environment of Educational Policy-Making in Large Cities*. Paper presented at the American Political Science Assoc. Convention, Sept. 1966.
- Campbell, Alan K. Who Governs the Schools? *Saturday Review* 64: Dec. 21, 1968.
- Campbell, Alan K. and Sacks, Seymour. *Metropolitan America: Fiscal Patterns and Governmental Systems*. New York: Free Press, 1967.
- Campbell, Roald F. et al. *Education and Urban Renaissance*. New York: John Wiley and Sons, 1969.
- Carter, Richard F. *Voters and Their Schools*. Stanford, Calif.: Institute for Communication Research, Stanford Univ., 1960.
- Carter, Richard F. and Sutthoff, John. *Communities and Their Schools*. Stanford, Calif.: Institute for Communication Research, Stanford Univ. 1960.
- Colton, David. *State Power and Local Decision-Making in Education*. Paper presented to the American Educational Research Assoc., Feb. 1969.
- Crain, Robert. *The Politics of School Desegregation*. Garden City, N.Y.: Doubleday, 1969.
- Crain, Robert L. and Street, David. School Desegregation and School Decision-Making. *Urban Affairs Quarterly* 2: 64-82; Sept. 1966.
- Decker, Michael. Representing the Educational Policy-Making Process: A Problem-Solving Approach. *Education and Urban Society* 1: 265-88; 1969.
- Dye, Thomas R. Government Structure, Urban Environment, and Educational Policy. *Midwest Journal of Political Science* 11: 353-80; 1967.
- Easton, David. *A Framework for Political Analysis*. Englewood Cliffs, N.J.: Prentice-Hall, 1965a.
- Easton, David. *A Systems Analysis of Political Life*. New York: John Wiley and Sons, 1965b.
- Elliot, Thomas H. Toward an Understanding of Public School Politics. *American Political Science Review* 52: 1032-51; 1959.
- Gittell, Marilyn and Hevesi, Alan G. *The Politics of Urban Education*. New York: Frederick A. Praeger, 1969.
- Gittell, Marilyn and Hollander, T. Edward. *Six Urban School Districts*. New York: Frederick A. Praeger, 1968.
- Gittell, Marilyn et al. *Investigation of Fiscally Independent and Dependent School Districts*. Research Project No. 3237. Washington, D.C.: U.S. Office of Education, 1968.
- Glaser, Barney and Strauss, Anselm. *The Discovery of Grounded Theory*. Chicago: Aldine, 1967.
- Goldhammer, Keith and Farnar, Frank. *The Jackson County Story*. Eugene, Ore.: Center for Advanced Study of Educational Administration, 1964.
- Goldhammer, Keith and Pellegrin, Roland J. *Jackson County Revisited*. Eugene, Ore.: Center for Advanced Study of Educational Administration, 1968.
- Greenstein, Fred. *Children and Politics*. New Haven, Conn.: Yale Univ. Press, 1965.
- Gross, Neal et al. *Explorations in Role Analysis: Studies of the School Superintendency Role*. New York: John Wiley and Sons, 1958.
- Gulick, Luther and Urwick, L. *Papers on the Science of Administration*. New York: Institute of Public Administration, Columbia Univ., 1937.
- Iannaccone, Laurence. *Politics in Education*. New York: The Center for Applied Research in Education, 1967.

- James, H. Thomas et al. *Determinants of Educational Expenditures in Large Cities of the United States*. Stanford, Calif.: School of Education, Stanford Univ., 1963.
- Kimbrough, Ralph B. *Political Power and Educational Decision-Making*. Chicago: Rand McNally, 1964.
- Kirst, Michael W. (editor). *The Politics of Education at Local, State, and Federal Levels*. Berkeley, Calif.: McCutchan, in press.
- Landau, Martin. On the Use of Functional Analysis in American Political Science. *Social Research* 35: 48-75; 1968.
- Lasswell, Harold and Lerner, David (editors). *The Policy Sciences*. Stanford, Calif.: Stanford Univ. Press, 1951.
- Lindblom, Charles E. *The Policy-Making Process*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.
- Machlup, Fritz. *The Production and Distribution of Knowledge in the United States*. Princeton, N.J.: Princeton Univ. Press, 1962.
- Martin, Roscoe C. *Government and the Suburban School*. Syracuse, N.Y.: Syracuse Univ. Press, 1962.
- Masters, Nicholas et al. *State Politics and Public Schools*. New York: Knopf, 1964.
- Munger, Frank J. and Fenno, Richard F., Jr. *National Politics and Federal Aid to Education*. Syracuse, N.Y.: Syracuse Univ. Press, 1962.
- Orfield, Gary. *The Reconstruction of Southern Education*. New York: John Wiley and Sons, 1969.
- Peabody, Robert L. and Rourke, Frances E. Public Bureaucracies. *Handbook of Organization*. (Edited by James G. March.) Chicago: Rand McNally, 1965. Pp. 802-37.
- Pois, Joseph. *The School Board Crisis: A Chicago Case Study*. Chicago: Educational Methods, 1964.
- Rogers, David. *110 Livingston Street*. New York: Random House, 1969.
- Rosenthal, Alan. *Governing Education*. Garden City, N.Y.: Anchor Books, 1969a.
- Rosenthal, Alan. *Pedagogues and Power: Teacher Groups in School Politics*. Syracuse, N.Y.: Syracuse Univ. Press, 1969b.
- Ross, Donald H. *Administration for Adaptability*. New York: Teachers College, Columbia Univ., 1951.
- Sayre, Wallace S. The Politics of Education. *Teachers College Record* 65: 178-83; 1963.
- Schoettle, Enid Bok Curtis. The State of the Art in Policy Studies. *The Study of Policy Formation*. (Edited by Raymond A. Bauer and Kenneth J. Gargen.) New York: The Free Press, 1968.
- Schreiber, Daniel (editor). *The School Dropout*. Washington, D.C.: National Education Assoc., 1964.
- Schutz, Richard E. Methodological Issues in Curriculum Research. *Review of Educational Research* 39: 364; 1969.
- Sharkansky, Ira and Hofferbert, Richard I. Dimensions of State Politics, Economics, and Public Policy. *American Political Science Review* 63: 867-79; 1969.
- Stroufe, Gerald. *Political Systems Analysis in Educational Administration: Can the Emperor be Clothed?* Paper presented at the American Educational Research Assoc., Feb. 1969.
- Sundquist, James. *Politics and Policy*. Washington, D.C.: the Brookings Institution, 1968.
- Truman, David B. Delusion and Regeneration: The Quest for a Discipline. *American Political Science Review* 59: 865-73; 1965.
- Tyler, Ralph W. The Field of Educational Research. *The Training and Nurture of Educational Researchers*. (Edited by Egon C. Guba and Stanley Elam.) Bloomington, Ind.: Phi Delta Kappa, 1965. Chapter 1.
- Usdan, Michael et al. *Education and State Politics*. New York: Teachers College Press, Columbia Univ., 1969.
- Zeigler, Harmon and Johnson, Karl. Educational Innovation and Politico-Economic Systems. *Education and Urban Society* 1: 161-76; 1969.

## AUTHORS

**MICHAEL W. KIRST** *Address:* Stanford University *Title:* Assistant Professor of Education and Business Administration *Age:* 30 *Degrees:* B.A., Dartmouth; M.P.A. and Ph.D., Harvard Univ. *Specialization:* Politics of Education; Evaluation of Federal and State Grant Programs; Use of Business Analytical Techniques for Educational Administration.

**EDITH K. MOSHER** *Address:* University of Virginia *Title:* Assistant Professor *Degrees:* B.A. and M.A., Texas Women's Univ.; Ph.D., Univ. of Calif., Berkeley *Specialization:* Educational Administration; Politics of Education; National Educational Planning; Organizational Theory.

## 7: ECONOMICS OF EDUCATION

MARY JEAN BOWMAN\*

*University of Chicago*

To discuss methodology in the economics of education one must say something about the nature of economics and how economists think. This is not to imply that all economists think alike, or even that the same economist thinks in the same way in differing contexts. Over the past decade some of the sharpest differences among men calling themselves economists have been clearly displayed in what has sometimes been labeled the "battle of the approaches"—approaches to educational planning and implied interpretation of how labor markets operate. However, starkly simplified initial positions on both sides are being re-interpreted and redefined, misunderstandings are being cleared up, and empirical evidence has been accumulating. More important, whatever the common and the disparate aspects of economic thought, two things must be made clear. Economics is a behavioral science and, at the same time, it is a branch of "moral philosophy." Systematic analysis of philosophical and ethical implications of the functioning of economic systems and of changes in those systems is quite as important a part of the economist's heritage as his more strictly behavioral or positive economics.

In both its behavioral and its normative aspects, economics is concerned with the quantitative analysis of interaction processes in partial or total *systems*. And in both of these aspects economics is based on the analysis of choices: economics could be defined as the science of *decision-making* regardless of the substantive content of the decision. Being rooted in a deliberately sparse but basic decision theory at the micro level and being concerned with the societal effects of interactive adjustments (a concern that is the reason for the micro-decision theory), economics displays a number of overlapping characteristics that are highly relevant to its methodology:

(1) Deductive mathematical formulations have long played a central role in economics, with the result that even when they are not making formal use of mathematics, economists use mathematical concepts. They easily extend their intuitions to an  $n$ -dimensional geometry, and differential

---

\* Dr. C. A. Anderson, University of Chicago, served as consultant to Dr. Bowman in the preparation of this chapter.



calculus serves naturally as a tool for simplifying micro-decision theory and the analysis of resource allocation in total societies.

(2) Economists make extensive use of equilibrium notions in the analysis of interaction systems, even when they focus on analysis of deviations from equilibrium or on fluctuations around it. Equilibrium concepts underlie important distinctions between short-run and long-run adjustments that are critical in any empirical analysis—for example, in understanding the interplay between educational developments and labor-market adjustments. Equilibrium concepts underlie the use of aggregate production functions in econometric studies of the sources of economic growth and the contributions of education to it. More broadly, they underlie a large part of econometrics, including use of simultaneous-equation techniques for the resolution of identification problems.

(3) Basically inseparable from (1) and (2), but calling for explicit mention, is the fact that economists make extensive empirical and theoretical use of the notion of incremental changes; in other words, marginal analysis is basic to economic theory. Often the marginal analysis may remain implicit because a writer takes it for granted and assumes his readers will do the same. But there are also some curious situations in which marginal analysis is simultaneously assumed by implication, yet overtly denied. One methodological consequence of the focus on marginal effects rather than averages is the economist's stress on regression coefficients rather than on proportions of variance explained. It is no accident that economists more often use the metric than the standardized forms of regression coefficients. They are much more concerned with "structural" than with "statistical" regression models, to use the distinction drawn by Kendall and Stuart (1963). Also, economists are especially fond of log-linear equations, partly because underlying distributions (especially of incomes) more nearly approach a log-normal form, but equally because of the importance in economic analysis of the notions of "elasticities" of response (and of elasticities of substitution in combinations of productive factors). The metric regression coefficients in a log-linear equation automatically give the elasticities of the dependent variable with respect to the independent variable.

(4) An important part of economics deals with maximization and minimization problems, or with optimization (or, more often empirically, with suboptimization). Particularly important for the economics of education have been two modern developments in applied research with such orientations: linear programming and the application of capital theory to benefit/cost assessments of educational investments.

(5) Occupying a conspicuous place in the discipline of economics and in political councils is national-income accounting. Behind modern developments in this sphere is a heritage of two centuries of concern with "national

wealth" and its measurement—two centuries of pragmatic compromise as well as ingenious invention. Techniques have become increasingly sophisticated with improvements in methods of index number construction and with the elaboration of economics input-output matrices. But none of these developments can get around the facts that (a) each of the items that together make up the national income, or the output of an economic sector, must be assigned a value (priced) to allow its direct comparison with other items, and that (b) the summing or aggregation of those values is theoretically legitimate only under special assumptions. Interpretations of national income accounts depend on how the pricing process and the nature of price and production relationships are conceived; so do ways in which national income accounts are used. Oversimplifying, two major contrasting points of view may be distinguished: the fixed-coefficient and the neo-classical. The fixed-coefficient view is associated with the elaboration of input-output matrices. It starts from an assumption of rigid production functions, or zero elasticities of substitution. "Manpower planning" in its most conventional form was a child of fixed-coefficient accountancy. The neo-classical view starts from assumptions of greater flexibility in production processes. Most empirical models in the neo-classical mode stipulate constant elasticities of substitution among production factors ("CES production functions"), though they differ in whether any further constraints are imposed by the model, and if so in the nature of those constraints. The researchers who attempted to explain aggregate contributions of education to growth in national income started from the neo-classical conceptualization and interpretation of economic relationships and the national income. That two such different starting points have been used has obvious import both methodologically and for substantive results.

(6) Finally, in the very nature of his subject and of its problems, an economist confronts the need to consider how the satisfaction of some ends for some people may affect the possibility or level of attainment of other ends (whether by the same or by other people). While this is part of behavioral or positive economics, it also goes back to moral philosophy and to the whole complex of questions concerning equity, opportunity, freedom, and "welfare" generally. It is decidedly bad economics, for example, to argue that people are behaving in a manner contrary to social welfare when they choose to go into lower paying ("less productive") jobs simply because they like that kind of work. Those who argue in such a way, though they seldom put it that baldly, draw a totally invalid distinction between so-called "consumer preferences" (expressions of which in the markets they do not challenge) and the realization of job preferences (which they place on a different and lower footing). I would put the job preferences first and justify that priority objectively by considering time allocations in the individuals' choices. Though further reference to the

philosophical heritage of economics is only incidental in this chapter, it would be a serious mistake to underrate the importance of analytical philosophy in the study of economic systems and in applications of economics to public (including educational) policy-making. For a good statement on this subject see Boulding (1969).

In this paper I first discuss uses of national-income accounting and aggregate production functions in the analysis of human-capital formation over time, and of its contributions to economic growth. A second section is concerned with empirical applications of capital theory to analysis of investments in human beings. Third, I present a few applications of multiple-equation techniques to the resolution of identification problems in the economics of education. The fourth section is focused on recent attempts at a dynamic analysis of the diffusion of schooling and of the ways in which schooling contributes to productivity. In the last section, I comment briefly about relationships among economic models for educational planning and the present state of the art.

### *Longitudinal Analysis of Human Capital in the Aggregate*

The work of three men encompasses most of the major variants in the methodology of longitudinal studies of human-capital formation over time and of contributions of education to growth in national income: T.W. Schultz (1960, 1961), Edward F. Denison (1962, 1967), and Zvi Griliches (1964, 1968). Though I can see relatively little direct transferability of these methods to other sorts of educational research (with some exceptions noted below), perspective on the methodologies behind these aggregative assessments is needed if only because findings of such studies have been so widely cited.

Schultz's major contribution was in opening up this subject by daring to make a first set of estimates, and especially in his treatment of *foregone earnings*. Denison described himself as a "standardizer," essentially applying index number techniques to analysis of aggregate inputs and of their relation to national income (final output). Griliches worked with aggregate production functions and has given the most attention to proper specification of variables and to "specification bias." I consider selected methodological features in the work of each of these men, with no pretense to a balanced summarization of their contributions.

#### *Schultz and Foregone Earnings*

Schultz's pioneering role in the economics of education is well recognized, but he succeeded more by the acuity of his intuitive insight, his persistence, and especially his genius as a catalyst than by his techniques.



Nevertheless, in his seminal work (Schultz 1960, 1961) on human-capital formation and on the contributions of education to economic growth, Schultz made a major methodological contribution; he applied the basic concept of opportunity cost to take account of foregone earnings of pupils and included the estimates of such earnings as part of aggregate human-capital formation at cost values. (Foregone earnings are those foregone when a pupil continues in school instead of taking a fulltime job.) Although there have been many arguments over Schultz's procedure, both in concept and in empirical application, the basic idea of counting student time as a resource input to the formation of human capital is little disputed today. More important, in my judgment, than its aggregative applications, this work by Schultz formed the starting point for quite different developments in micro-decision theory and its empirical applications, developments which I present in the next section.

One of the most pervasive arguments concerning empirical procedures in the estimation of foregone earnings as a measure of student inputs into human-capital formation relates to corrections for unemployment. Critics have commonly asked: "What if all the students were thrown on the labor market at once?" Yet the same critics never ask: "What if all the teachers were thrown on the labor market at once?" Two methodological-conceptual fallacies are involved: 1) overlooking the fact that foregone earnings are *like all prices* (including all the estimates that enter into national income accounts) in measuring the value of a good or service *in a given economic context only* and 2) confusing which measures are proxies for which underlying variables (concepts) in a particular problem. In investigating resource allocation, which requires comparison of one alternative with another, "foregone earnings of students" measures the alternative properly, whereas salaries of teachers are at best proxies for what the teacher could earn in some other occupation. But, if the purpose is to assess the value of the resources in fact used in the schools, asking whether those resources (teachers' time, students' time, etc.) would be used somewhere else if they were not used for education becomes totally irrelevant. In totalling output for the construction industry, no one would suggest that the estimator should take into account that if the bricklayers were laid off they might not get another job easily. The "what is put in" view in this sense would treat the foregone earnings of students as the proxy measure for value of their inputs; if they indeed are students, one does not ask in this context, what would happen if they were not—nor does he properly make any adjustment for unemployment rates. For discussions of this issue see Bowman (1966).

Two defects appear in Schultz's (1961) estimates of the contribution of education to growth (see Bowman, 1964). First, it is only because of a computation error that his estimates came out as they did, and no em-



pirical weight can be attached to them; but this is not methodologically important. The other, and methodologically important, point is that by taking a round-about route to arrive at what education contributes (bringing in human-capital costs only to multiply them out again), Schultz made the mistake of applying an "internal rate of return" instead of a rental-value ratio. That is, instead of applying an ongoing estimate of what a unit of human capital would produce through time, he discounted that return back to the date at which the investment was made.\* This slip reflected that he was interested in the human-investment idea and carried it over inappropriately to the quite different problem of measuring aggregate contributions of human capital to growth in national income. Denison and Griliches avoided this error, though they owe much to Schultz's insights.

### *Denison and How Sources of Growth Differ*

Denison has long been an outstanding contributor to analytically-based measurement of aggregate economic inputs and the analysis of changes in those inputs over time. The basic purpose of his last two books (on the United States, 1962, and on Europe, 1967) was to find as full an explanation as possible for the "unexplained residuals"—that part of growth in national income over and above what could be explained by the aggregate inputs of (undifferentiated) labor and (non-human) capital as conventionally measured. He used index-number techniques to adjust for changes in labor quality; his weights were derived from data on earning differentials by education (i.e., the "rental values" of the education embodied in members of the labor force). These rental values were adjusted downward by forty per cent to allow for the effects of ability differentials associated with schooling differentials (but presumably independent of schooling) upon the observed patterns in earnings. Similarly, he made adjustments for changes in hours worked and in age and sex composition of the labor force. These adjustments for labor quality implicitly take into account the post-school investments in human-resource formation on the job, in addition to effects of schooling *per se*.

Denison also attempted estimates of effects of scale on productivity, the influence of improvements in resource allocation, etc. Nevertheless, in all cases a substantial unexplained residual (which he termed "advancement in knowledge") remained. Included in that residual were any changes in the quality of education or of physical capital that could not be accounted for by increased inputs into education or physical-capital formation. In other words, in the idiom of this field, Denison's advancement in know-

\* This again would have led to a seriously erroneous estimate of the contribution of human capital were it not for the compensating effect of relationships between age and earning differentials associated with schooling and of the distribution of education by age among members of the labor force.

ledge is treated as an increase in total productivity. Note that in all such studies what is called "increase in total productivity" is going to depend on which factors the analyst deliberately or inadvertently fails to specify.

Denison's findings were dramatic—not only for "Why Growth Rates Differ" (Denison, 1967), but also for *how sources of growth differ*. With respect to education in particular, he estimated a per annum contribution to growth in per capita incomes over the period 1950-1962 of .49 percentage points in the United States against an average of only .23 points for Northwestern Europe. Because per annum growth rates over-all were lower in the United States, the percentage of growth in per capita incomes attributed to education in the United States was almost four times the percentage credited to education in Northwestern Europe (25% against 6%). It is interesting to note also that the mean per capita income level of the Northwestern European countries in 1960 just matched that of the United States in 1925, despite the fact that the educational attainments of the labor force in this country as of 1925 were considerably lower than those of Europe today. What does all this mean and what methodological factors are involved?

First, this, like most aggregative assessments of contributions of education, is really a description of the combined effect of (a) prevailing differentials in earning associated with differentials in schooling and of (b) changes in the educational composition of the labor force. On the basis of (a) alone, one would expect a higher contribution of schooling to income growth in Europe than in the United States, since relative differentials in earnings were higher there. However, mass education in the United States (which began early enough to extend its effects by 1960 over a large proportion of the labor force even at middle age or older) expanded enough to sweep education into a front position among factors explaining growth. In other words, in this type of analysis the estimates of what education explains are predetermined by the estimates of measured inputs of education at any given constant pattern of differentials in earnings. There is no independent test of how those inputs affect outputs over time. They could even *over-explain* observed outputs; this occurred in an application of essentially this technique to Russia for the period between World Wars I and II (Dewitt, 1962).

Second, Denison's work, along with the work of several others, shows that relative differentials in earning have remained remarkably stable in the United States despite the massive diffusion of schooling. The question of why this has been the case has attracted increasing attention among economists in the past year or so. But this is to remark also that there is an inherent fallacy in even asking "how much has education contributed to growth?" since that question can make sense only for a given Gestalt of other resources and for a given economic system. Aggregate estimates

of single-factor contributions to growth are a bit like maps of the United States drawn in the seventeenth century; they have some recognizable resemblance to reality, but if a traveler took them literally he would be in for some very big surprises.

### *Griliches, Specification Bias and "Total Inputs"*

Unlike Denison, Griliches has fitted equations in time series analyses of determinants of national income. Like Denison, he has attempted to specify more fully the inputs into production, seeking thereby to provide a fuller explanation of growth.

But Griliches has pushed much further than Denison in his insistence on full specification and the inclusion of all "inputs." This is probably the aspect of his published work that would be of greatest interest to researchers in education. Ideally, he wants to include specification of those improvements in the "quality" of capital and of education that Denison left in his residual category "advancement in knowledge." Despite his seeming absorption of the "residual" in his latest attempt (Jorgenson and Griliches, 1967) to account for national income growth, how far he has succeeded in specification of changes in quality of capital, for example, must be seriously questioned. The important point for this chapter is the conceptual position and its methodological implications—the position that if one could adequately measure inputs, he would fully explain growth (and not over explain it, since negative inputs would also be fully specified). Along with raw cotton and embodied schooling (measured in unchanged schooling-quality units), Denison's advancements in knowledge, in one form or another, are to be treated as "inputs," all essentially in the same way. This point of view could be carried over at a quite different level, to study of inputs and outputs of school systems or of individual schools. The major difficulties are in finding ways to measure the heretofore ignored inputs without circularity and to identify what part of the incremental earning power empirically associated with schooling, for example, is indeed attributable to it. As an outgrowth of his interests in proper and full specification of variables and of their effects, Griliches has recently become very much interested in the handling of "ability" in econometric models analyzing economic growth.

### *Human Investment Decision and Its Empirical Assessment*

Methodology in the economics of education can hardly be discussed apart from the applications of capital theory to analysis of investment in human beings. This theoretical construct both sets out the empirical parameters of interest to many economic researchers and undergirds or gives

meaning to a large part of the behavioral analysis relating to education in school and at work—including the relationships among education, labor market structures and productivity. The basic human-investment decision model constitutes the prototype, also, for benefit/cost analysis in its societal, public policy applications.

### *The Skeleton Model*

In its bare bones, the decision model has just two basic parts: (1) the comparison of a man's expected life-income streams (including negative components) in pursuing one course with the stream expected if he were to choose the best alternative to that course and (2) a way of adjusting for the timing of income or earnings so that streams with different shapes through time can be compared with each other. Suppose a youth is considering whether to go to college. If he goes directly into the job market when he graduates from high school, he will start earning at, let us say, age 18, and he need make no further direct monetary outlay on schooling. Furthermore, he will be acquiring experience at work that increases his earning power. However, a time will come when his earnings are less than they would have been had he chosen to go through college.

If he chooses to attend college for four years (in this simplified model), he will earn nothing and will have direct costs of tuition, books, etc. that he would not incur if he were working;\* his income for these four years is negative to the amount of his direct outlays. It may be that when he first starts to work he still will not make as much as he might have at age 22 had he been working throughout the interim. But eventually and usually fairly soon, his earnings will be greater than if he had gone directly from high school to work. How does one then proceed to analyze this decision situation?

There are two streams of life income; at some point the college stream and the high school stream cross each other. There are several different ways to proceed. One would be to take each stream and consider, from the perspective of the initial decision point, how much that stream was worth. But one can not just add the incomes for every year without paying attention to whether they came early or late in a life sequence; most people would prefer \$1,000 now to a guaranteed \$1,000 ten years hence. Even if he were to postpone spending it, the recipient of \$1,000 now would be better off since he could invest it and have more money later. So the individual may apply an "external" (or "reservation") discount rate to each of the streams and compare them. (For a clear discussion of "external discount rates as "reservation" rates see Harvey, 1967, chapter 2.) If the college stream is worth more at this discount rate, he will "decide" to make the

\* He would have to cover subsistence in any case; this is not a cost of schooling.



*investment*; that is, he will give something up now in return for a claim on additional income in the future. Another way of arriving at the same conclusion (again by a present-value method) would be to subtract the high school from the college stream and then discount to get the present value of the *differential stream*. (The first part of that differential stream, at least during the college years, will be negative; it is made up of the foregone earnings that would be made by the high school graduate plus direct outlays on college expenses.)

The *internal rate of return* in this example would simply be that discount rate which would make the college and high school streams of *equal present value*. The internal rate of return is thus a very convenient way of comparing investments. There is a considerable technical literature on the uses and limitations of these "statistics" (for that is what they are) and on their interpretation in economic terms. (See, for example, Hirschleifer, 1958, and Feldstein and Fleming, 1964.)

### *Analysis of Investments in Training and Learning at Work*

In what can only be regarded as a *tour-de-force*, Becker (1962, 1964) and Mincer (1962) extended rate-of-return analysis and "opportunity cost" in the analysis of investment in human beings to on-the-job training or learning and to several other applications. Under certain assumptions concerning labor market structures, this analysis allows one to estimate the amounts individuals invest in themselves when they choose jobs in which immediate earnings are less but earning power increases with experience. Other labor market situations lead to an on-the-job human investment in which the employer incurs the costs and from which he expects to reap the returns. The analytical scheme developed by Becker and Mincer has opened up a range of important questions, and provided a meaningful framework for research on them. It has stimulated explorations of the nature of learning functions: learning to learn versus to earn, and the timing and interdependence of these kinds of learning. Ideas from rate-of-return analysis and opportunity costing combined with distinctions in pertinent characteristics among labor market structures have stimulated new research into hidden investments by business enterprises in the formation of human capital, the conditions that encourage such investment, and approximate measurements of its extent. Application of these methods permitted a fresh empirical perspective on the workings of discrimination in labor markets *vis-a-vis* minority racial groups and on the nature of labor market disadvantages (not necessarily discriminatory) of women. My current explorations of labor market structures in Japan and of the labor market as a capital market stem from this same theoretical-methodological base.

Some of the empirical work that emanated from these developments makes use of one or another sort of multivariate statistical analysis along with present-value and rate-of-return accounting. Regressions, usually using sets of dummy independent variables, are used, for example, in attempts to purify the estimates of income differentials attributable to schooling—a necessary statistic for proper interpretation of rate-of-return estimates whether for analysis of private or of social decision making. In a few cases there has been more extensive use of statistical methods; the ongoing econometric research into formation and acquisition of human capital as a business asset (under the leadership of Lester Telser) and my current work on Japan are examples. The key relevant variables in this work are quite other than might have been stressed ten years ago.

### *Social Evaluations and Benefit Cost Assessments*

It has sometimes been assumed, and this can be found easily in the writings of some economists, that rate-of-return analysis rests upon the assumption that competitive-equilibrium prevails. This is a basic misunderstanding of the method and the scope or limitations of its applications. In fact, rate-of-return analysis constitutes part of the apparatus with which the economist searches for monopoly, non-monetary preferences, concealed non-homogeneities, etc. Interest in such evaluation has been declining recently in the face of pressures on and enticing bribes offered to economists to engage in investigations directed toward certain public policies.

In education, as in a wide range of public goods (water resource development, highways, etc.), there has been a rapidly accelerating spread of benefit/cost studies. These can be extremely complex and elusive in execution. Investigators run into one seemingly immeasurable but important variable or set of variables after another, and the problems of proper specification and identification of relationships are myriad. Nevertheless, the core of social benefit/cost analysis and its skeletal methodology are essentially the model presented in the first part of this section. In education, that model is usually applied with only minor modification. For the schooling period, foregone earnings of college students still count as a cost (these are socially foregone products), but tuition payments will be excluded; school expenditures on equipment, services of teachers and other staff, etc. are included since these resources are diverted from other areas. Whereas for the individual the relevant income streams are taken after payment of income taxes, from the societal point of view they should be taken before taxes.

All this is straightforward enough. However, even in attempts to estimate "true" private rates of return many problems are encountered—for example, the translation of data from age-income cross-sections of the

population at a given time into appropriate longitudinal estimates for a particular cohort. In social benefit/cost accounting there are more serious complications, especially those associated with scale and "spill-over" effects, that can be disregarded in analysis of individual decision. Moreover, any social decision has distributional effects that are irrelevant at the individual level. Conventional benefit/cost analysis is an indispensable but insufficient tool for educational planning.

### *Applications of Multiple-Equation Techniques to Identification Problems*

Situations in which the coefficients of an equation can not be identified uniquely are extremely common in all the social sciences, though often they go unrecognized. Economics is distinctive on this score only in that it has been favored from its origin by the easy availability of money measures, and hence, more than the other social sciences, economics uses quantitative models of interactions in partial and in total systems. (Economics has been also especially concerned with analysis of time series.) It is no accident that discussions of multicollinearity go back 35 years or more in economics (see Frisch, 1934) or that the distinct but partially overlapping problem of identification should have received special attention from econometricians.\*

Although the use of recursive equation systems and of simultaneous equations are techniques that have thus far been employed in the economics of education to a limited extent only, there can be little doubt that their use in that field will spread in the near future. Moreover, these techniques have potential applicability to a wide range of problems in educational research that would not normally be regarded as "economic."\*\* For these

---

\* For discussions of the identification problem, see Hood and Koopmans (1953). A particularly obvious example of ambiguity of coefficients (their underdetermination) is provided by the problem of empirical applications of demand-supply analysis. Quantity demanded is a declining function of price, quantity supplied is a rising function. Application of a simple equation relating amount exchanged to price gives coefficients that are highly ambiguous. This is the classic introduction to simultaneous equation methods and their role in empirical economics. But equally important are other multiple-equation techniques that have more in common with the path-analysis techniques developed by statisticians in biology and adapted from them, more recently, in sociology.

\*\*Of special interest to educationists must be the big "opportunity study" directed by Coleman (1966). That study stimulated active methodological and substantive challenge by and discussion among economists. See, for example, the series of comments and responses in the *Journal of Human Resources* (Bowles and Levin, 1968). Methodologically, the argument was concerned almost entirely with step regressions and the implications of orders of entry of variables in such analyses. However, I expect that multiple-equation techniques will soon be applied to analysis of a number of problems concerning interactive effects of school and home variables on measured achievement. Proper identification of production functions is indeed necessary for sound cost-effectiveness or benefit/cost analysis in educational decision making. For an extended discussion of some of the methodological problems see Bowles (1968).



reasons I present three examples here, each of which illustrates some distinctive method or mode of application.

The pertinent problems in each of the three examples can be stated briefly: The first, which is the simplest, is identification of the extent to which and the paths by which "education" affects earnings differentials. The second is identification of manpower supply and demand forces and their interactive impact on the structure of the labor force (using international comparisons). The third is the disentangling of effects of interdependence between per capita income and public expenditures on education, with particular attention to the nature and extent of "single-equation bias" in estimating elasticities in each direction of influence. For easy reference I label these studies respectively the *education-earning paths* model, the *manpower-balancing* model, and the *simultaneous income-education* model. The second is as yet only a suggestion or a hope; it is being explored in the Paris offices of the Organization for Economic Cooperation and Development (OECD). The first and third models have been completed but have not been published. The first two are recursive models; that is, each equation is independent of the others. The third is an interdependent or "simultaneous" equation system.

### *The Education-Earnings Paths Model*

Paralleling the developments in path analysis in biology and in sociology has been the development, lead by Herman Wold, of multi-stage regressions in empirical economics. A simple but interesting application of this technique to problems in the economics of education is provided in a paper by Weiss (1969). I confine my remarks here to the recursive equation system, which he used with whites only—except to note that in specification of the education variable Weiss developed a quality-controlled measure of education, using mean achievement scores by locality to adjust the raw years of schooling. These adjusted values (assigned to individuals according to their location and years of schooling) are designated by the letter  $H$ . The first stage equation then takes occupation (using a standard occupational status scale) as the dependent variable, with  $H$  as an independent variable,  $b_1$  being the regression coefficient on  $H$ . A second equation treats weekly earnings ( $P$ ) as the dependent variable, putting  $H$  (with a coefficient  $b_2$  and occupation ( $O$ ), with a coefficient I labeled  $c_2$ , along with other control variables, on the right-hand side. A third equation does the same with weeks worked ( $W$ ) as the dependent variable, and the coefficients  $b_3$  and  $c_3$  on  $H$  and  $O$  respectively. Finally, earnings ( $Y$ ) are simply the product of  $W$  and  $P$ . Taking first differences, rearranging, and substituting, Weiss arrived at an equation that sorts out the direct and indirect effects of education on earnings as follows:



Effects on Average  
Earnings  
Within      Via  
Occup.      Occup.

Effects on Employment  
Within      Via  
Occup.      Occup.

Interaction between  
Employment and  
Weekly Wage

$$\frac{\Delta Y}{\Delta H} = Wb_2 + Wb_1c_2 + Pb_1 + Pb_1c_3 + (b_2 + b_1c_1)(b_3 + b_1c_1) \Delta H$$

These equations were estimated separately for successive age classes for white males. The interpretation of the results require caution and theoretical and intuitive insights that go beyond the statistical model. It is interesting to note that Weiss found those of the effects of education on earnings that operated by way of effects on occupations to rise the most rapidly with age. It is clear, even though the recursive equations were not applied to the data for non-whites, that no such pattern would be found among them. While these findings are hardly surprising, they have interesting implications (understandably missed by Weiss) when related to sociological analysis of discrimination taken in conjunction with the recent developments in analysis of on-the-job training and the shapes of life-income streams to which reference was made in section 2.

### *The Manpower-Balancing Model*

Internationally, no field of work linked to the economics of education has attracted more effort from economists and accounting statisticians than the so-called "manpower planning" and the forecasting of "manpower requirements."

The so-called "manpower approach" was built initially on a simplified model of the economy that assumed demand domination of the labor market, with virtually zero elasticities of substitution among productive factors. It allowed no place for independent influences of manpower supplies on labor force composition at any given level of national per-capita income—which is commonly treated as synonymous with "technological level." The OECD has been very actively involved in this type of endeavor; it has been a participant in manpower-planning efforts and in the compilation of a detailed compendium of data on the occupational, industrial and educational composition of the labor force, both in OECD countries and in other nations for which information could be obtained (OECD, 1969). The OECD has also invited and itself undertaken critical appraisal of these endeavors and of the presuppositions upon which manpower-requirements forecasting was based. All this has led to a quite fundamental reassessment. First came the extremely important *Technical Evaluation of the Mediterranean Regional Project* prepared by Hollister (1966). Hollister especially stressed the importance of considering independent supply influences on labor market adjustments and the relationships between labor force composition and national income. Following that study,

OECD undertook an in-house analysis of the accumulated data in its statistical compendium, and in July 1969 specialists were brought together in Paris to discuss that work and make suggestions for its improvement. The problems of the identification in interdependencies in manpower demands and supplies was central in the discussions, and several partial attacks on this problem were suggested. The most interesting and complete model was put forward by Malinvaud.

Malinvaud's model treats the educational mix in the labor force at large as a supply constraint in any given "present," determined by prior actions. He treats the occupational structure quite independently as a reflection of demand from the productive system. The proportions with each level of education among those in any given occupation are seen then as the result of balancing adjustments by which manpower demand and supplies are matched. The sets of independent demand and supply equations are as follows:

$$L_{jt}/L_t = f_j(n_t) + e_{jt} \text{ (the demand equations)}$$

$$L_{kt}/L_t = f_k(n^*) + e_{kt} \text{ (the supply equations)}$$

where:

$L_t$  is the number in the total labor force of country  $t$

$L_{jt}$  is the number in occupation  $j$  in country  $t$

$L_{kt}$  is the number with education  $k$  in country  $t$

$n$  is a set of variables determining current demands for labor in the various occupations.

$n^*$  is a set of variables (which could be the same as or different from variables in  $n$ ) determining the past schooling decisions of present members of the labor force.

Malinvaud, along with other participants in the conference, suggested a longer lag period on the supply equation (giving  $n^*$  an earlier dating).

The adjustment equations of Malinvaud's model began with assuming a "normal" structure of each occupation by education level, that structure being defined by a set of coefficients  $a_{jk}$  such that for each  $j$

$$\sum_k a_{jk} = 1 \quad (3)$$

The condition  $\sum_k a_{jk} = 1$  is of course a necessary condition for consistency

in a total system; that is, for each country  $t$  the sum over all  $k$  of the proportions of any  $L_j$  who are  $k$  must add to one. The "normal" demand for  $L_k$  following equation (3) then becomes:

$$\sum_j a_{jk} L_{jt} \quad (4)$$

The deviation of  $L_k$  from the "normal" number predicted by (4) would lead to a ratio  $L_{jkt}/L_{jt}$  in country  $t$  that deviated from the "normal" coefficient  $a_{jk}$ . Malinvaud suggested, finally, the equation in (linear, not logarithmic, form):

$$\frac{L_{jkt}}{L_{jt}} = a_{jk} + b_{jk} \left[ \frac{L_{kt}}{L_t} - \sum_j a_{jk} \frac{L_j}{L_t} \right] + e_{jkt}. \quad (5)$$

The sum over  $k$  of  $e_{jk}$  must be zero, and the second term on the right hand side of this equation defines the expected deviation from the "normal"  $a_{jk}$ . That deviation will be the (bracketed) difference between the actual and the predicted proportions of the labor force who are in education category  $k$  multiplied by the coefficient  $b_{jk}$ , which is allowed to vary over both  $j$  and  $k$ .

It is important to recognize what such a model is and what it is not. The variables in equations (4) and (5) remained unspecified—that is, the reference was to variables already available to and used by OECD in its cross-national analysis. Though OECD used a variety of special indicators, it relied in the main on a measure of per capita income or, alternatively, a non-monetary index of economic or technological level of development. No generally applicable measures of physical capital were used.\* But these specifications could change without altering the main features of the model. Its most critical characteristics are in its recursive nature combined with its consistency constraints.

Equation (5) is clearly and unambiguously in the manpower-requirements tradition; it contains no wage rates and takes the *occupational structure* as entirely independent of conditions of educational supply. This is workable only with a very broad categorization of occupations; each occupation is only vaguely specified and encompasses a wide range in tasks and in relative pay. The adjustment equations give supply an influence on the organization and distribution of tasks and responsibilities, but only *within* the occupational rubrics defined by the "demand" side of the analysis; supply conditions are not permitted to affect the distribution of the labor force *among* occupations or, alternatively, among sectors.

Equation (4) is concerned with human investment decisions only in the most incidental fashion. If past per capita income, for example, is a good statistical predictor of present educational mix in the labor force as a whole, that is enough. How far individual decisions with respect to schooling are or are not responsive to expected rates of return on such investments is a matter that is necessarily excluded from consideration as long as the recursivity of the system is maintained. This could be given an economic

\* However, OECD used proxy measures designed to pick this up when working with particular economic sectors, and these could be introduced into equation (2) if the above system of equations were applied with the  $j$  standing for economic sector instead of "occupation."

interpretation if the model is viewed as essentially concerned with short-term adjustments; in such a context the present educational composition of the labor force is an exogenous variable—as is the independent demand structure. The model does not provide a dynamic analysis of effects of current adjustments on future manpower supplies.

These comments are not criticisms. To dynamize the model, making it a genuine development model that traces relationships through time, would require longitudinal series of data within countries that are not available. Moreover, the difficulties in making a complex dynamic model operational would be immense even were data readily available. This model, in other words, is a short-term adjustment model of considerable flexibility in application, but with no explicit base in the analysis of decision processes and with no consideration of the *longer-term interactions* between manpower demands and supplies.

### *A Simultaneous Income-Education Model*

The study of interdependence between public school expenditures and per capita incomes across the United States, by Tolley and Olson (1969), is almost the opposite of Malinvaud's model in several important respects. First, Tolley and Olson made use of simultaneous equations and their central concern was with the interdependency among the equations of the system. At the same time, they were particularly interested in assessment of the magnitude of "single-equation bias" in estimates of elasticities of educational expenditures with respect to incomes and vice versa. Second, the Tolley-Olson study was grounded in human-investment decision theory. And third, it was concerned not with consistency in predictions and the specification of constraints that will ensure that result, but rather with sorting out reciprocal effects of two variables when the two are mutually supportive in their effects—higher income leading to higher educational expenditure and in turn to higher income, and so on.

Before summarizing the Tolley and Olson paper, a word is in order concerning relevant prior work and the state of ignorance in the sorting out of aggregative effects of education on per capita incomes and vice versa across nations. This has been a widely recognized problem. But it is also a topic on which some of the most shocking methodological sins have been committed. It should go without saying, for example, that use of data on current school enrollment rates to "explain" a nation's current per capita income is entirely unjustified; yet there have been economists with international reputations who have based pronouncements concerning the "effects of education on incomes" on data on proportions of children currently attending school. Fallacies of using single factors in comparisons among widely differing societies aside, it would appear that the obvious answer would be to get the time sequence in the right order—if data re-



garding the educational composition of the adult population are not available, the appropriate child-enrollment rates are obviously those applicable when today's adults were of school age. Conversely, today's income levels are obviously appropriate when the dependent variable is today's enrollment rates or expenditures on schools. However, the proper timing and lagging of variables does not provide an escape from problems of serial correlation.

Ten years ago, when the first crude correlation analyses using cross-section data by nations to examine relationships between income and education were appearing, C. Arnold Anderson and I demonstrated the empirical importance of such serial correlations on the international scale, along with a discussion of the circularity of causal relationships between income and schooling. We showed that not only did incomes at a given time explain enrollment rates at that time better than adult educational attainments explained incomes; the nature of influences at work and their continuities were such that we actually obtained highest correlations by going to the extreme of using incomes of the 1930's to explain enrollment rates in the 1950's. (See Bowman and Anderson, 1963.) However good the data may be, interdependence between education and income in a mutually supportive spiral can produce very marked single-equation bias in regression coefficients, whether taking education as a function of income or income as a function of education.

Tolley and Olson concentrated on the interdependencies between income and public expenditures on education among the states of the United States, not on years or levels of schooling attained, though the latter also enters into their equations. They began by setting up behavioral equations to estimate relationships in each direction of causation separately and then simultaneously.

The equation treating per employee income ( $Y$ ) as a dependent variable specifies a log-linear production function, the independent variables of which include estimates of non-human capital ( $K$ ) and of human capital. Human capital is specified as a function of years of schooling ( $S$ ) and of quality of schooling. Quality in turn is a function of expenditures per pupil ( $E$ ). (The zero-order coefficient of determination between lagged (1929) and unlagged (1960) values of  $E$  was .75; that between lagged and unlagged incomes was .96.) Assuming exponential forms of the variables, the equation for income is then:

$$Y = a_1 + b_1E + c_1U + d_1D + e_1N + f_1S + g_1K + u_1 \quad (6)$$

where  $U$ ,  $D$ , and  $N$  are control variables to correct for local differences in costs of a particular "quality" of education. Given the log-linear form of the equation, the regression coefficients are all elasticity coefficients. By far the highest of these was for median years of schooling of the adult popu-

lation. Educational expenditure per pupil in average daily attendance ( $E$ ) was second in importance; using a lagged (earlier) expenditure variable gave very similar results, mainly because of the high correlation between lagged and unlagged versions of  $E$ .

Tolley and Olson approached the construction of an equation with educational expenditure per pupil as the dependent variable by first stipulating the conditions (derived from economic decision theory) under which one would expect equal expenditures per pupil in all states. Those conditions included: "(1) identical education cost curves, (2) no consumption value of education, (3) equal earnings for persons of the same education every where in the country, and (4) possibilities to borrow unlimited amounts at an interest rate which was the same in all States." Explanation of variance in expenditures among the States was sought by relaxing these conditions. The final equation to predict expenditures per pupil looks very like that explaining per capita income so far as the variables included are concerned; the only change, other than the exchange between expenditures and income as dependent variable, is the addition of the independent variable, pupil/population ratio, ( $P$ ). This equation reads:

$$E = a_2 + b_2Y + c_2U + d_2D + e_2N + f_2S + g_2K + h_2P + u_2. \quad (7)$$

To arrive at this seemingly simple formulation, the authors performed a complex and sophisticated theoretical-mathematical analysis, which is ingenious and in itself illustrative of an important methodology in economics. Unfortunately, it cannot be adequately described here.

Solving equations (6) and (7) for the two endogenous variables ( $E$  and  $Y$ ), they derive the "reduced-form equations"

$$Y = a_3 + c_3U + d_3D + e_3N + f_3S + g_3K + h_3P + u_3 \text{ and} \quad (8)$$

$$E = a_4 + c_4U + d_4D + e_4N + f_4S + g_4K + h_4P + u_4. \quad (9)$$

Since only one endogenous variable is included in each equation, the regression coefficients on the other variables are free of the biases in direct least squares estimated of (6) and (7). Each of the coefficients in equation (8) and (9) can be expressed in terms of coefficients in equations (6) and (7), which allows Tolley and Olson to distinguish the direct from the indirect effects of the various exogenous variables (indirect via their effects on the endogenous education and income variables and via interactions between the latter). However, to attain identification of the elasticity coefficient  $b_1$  (regressing income on educational expenditures) from the reduced-form equations, there must be one and only one independent variable that appears in equation (7) but not in equation (6), and conversely for  $b_2$ .

The condition for identification of  $b_1$  is met since the variable pupil/population ratio appears in equation (7) but not in equation (6). However,

a problem arises with respect to identification of  $b_2$ . This is discussed at length in the Tolley-Olson essay, which arrives (through what I regard as entirely reasonable assumptions) at a revised form of equation (7) that omits the independent variable ( $S$ ), which refers to quantity (average years) of schooling of the adults. Among their empirical results in estimation of elasticity coefficients, those shown in Table I are especially interesting.

TABLE I

SELECTED EDUCATION-INCOME ELASTICITY COEFFICIENTS:  
SINGLE-EQUATION AND SIMULTANEOUS-EQUATION  
ESTIMATES COMPARED

<i>Coefficients</i>	<i>Single Equation Estimates</i>	<i>Simultaneous Equation Estimates</i>
Elasticity of income with respect to:		
Educational expenditures ( $b_1$ )	.259	.066
Years of schooling per capita	.649	.852
Elasticity of education expenditures with respect to:		
Income from sources other than private non-human wealth ( $b_2$ )	.791	.869
Income from private non-human wealth	2.22	1.65
Pupil/population ratio	-.478	-.473

Substantive evaluation of these results would require a careful examination of details of the model that are not reproduced here. For example, are the exogenous variables properly specified in accordance with their place in the theoretical construct? Especially important in this respect are the proxies for non-human capital and for incomes from it. Serious questions may be raised about the fact that "years of schooling" was treated as an exogenous rather than as an endogenous variable. However, the authors experimented with a number of variations in specification of their independent variables, and they tried out alternative measures of  $K$ . They found that consistently the single equation bias for  $b_1$  was substantially greater than for  $b_2$ ; this is associated with the fact that the absolute value of  $b_2$  is much greater than the value of  $b_1$ , making  $E$  highly dependent on  $Y$  instead of independent, as assumed in single equation estimates of  $b_1$ . In all cases the simultaneous-equation estimate of elasticity of educational expenditures on income from private non-human wealth was considerably greater than on income from other sources.

## *Methodological Forays in the Analysis of Schooling, Communications, and Innovation*

Few economists would question the importance to a decision-maker of knowledge with respect to the basic parameters of his decision—though there might be considerable argument as to how far, under what circumstances, relevant knowledge is in fact available to youth or parents in making decisions concerning human investments or schooling. Neither would any economist deny that in order to understand how education contributes to productivity it would be helpful to identify, for example, spill-over effects in the diffusion of information by educated people (not merely teachers). And more clues are needed concerning just *how* schooling contributes to adaptive and innovative behavior. Studies of such questions are just beginning. The dearth of work relating education to communication and to innovation is hardly surprising, given the difficulties in these questions, whether conceptual or in empirical application. But there is some recent work of considerable interest.

In discussing this literature I shall focus upon (1) the integration of communication theory into analysis of the human-investment decision, and (2) a cluster of investigations concerned with identification of the innovation component of individual earnings differentials as related to schooling, and with the "spill-over" effects of schooling as related to schooled members of a farming population.

### *Communications and the Human Investment Decision*

As far as I am aware, the incorporation of communication factors in models of decisions about schooling has been confined to work of staff and students at the Comparative Education Center of the University of Chicago. With some partial exceptions in the work of Chaudhri (1968), the Chicago Center has also been doing the only studies that incorporate communications into analysis of post-school human resource development and utilization. In all of the Chicago studies the basic theoretical framework derives from a merging of economic decision theory with Torsten Hägerstrand's (1968) work on the diffusion of innovations. The economic model fits into and elaborates the "resistance" side of the Hägerstrand model, whereas his analysis has concentrated much more on the "information field" side, but both are necessary. Of these Chicago studies, the one closest to completion is Goldblatt's examination of determinants of enrollment rates (and related dependent variables) among children in the states of Mexico.\*

\* I refer here to research being done to follow up Goldblatt's (1968) dissertation. Her new work incorporates a number of important methodological improvements, though it is built upon the same theoretical structure as her dissertation.



The basic decision model specifies four parameters: (1) participation in *information fields* that bring knowledge of returns to schooling in the modern sectors to the people of an area; (2) the extent of local (within-the-State) access to preferred jobs and higher future earnings for those with more education; (3) opportunities for economically productive activity on the part of children if they drop out of school (or never attend); and (4) ability to give up potential earnings of children (and to meet direct costs) in the present in anticipation of future returns. This fourth parameter is important because of the severely limited capital markets for financing investments of individuals in themselves.

The model assumes that there are no significant differences among States in direct costs of providing schooling at primary levels except such differences as may exist between urban and rural areas; most of the analysis is for either urban or rural populations, treated separately. Goldblatt discarded a variable that had been urged by some economists looking at her work, as a proxy for a severe cost restraint: namely, proportions of schools lacking places beyond the third grade. The decisive reason for excluding this variable is that, whereas this is indeed a major constraint on the cost side from the point of view of an individual decision maker, it takes on quite another meaning with reference to the entire population of a state or province, becoming unquestionably more a reflection of differences among States in *demand* for school than in conditions of supply. Economists who have failed to appreciate this are displaying an unawareness of the ecological fallacy that is all too common; sociologists are less often trapped in that way.

The equations used by Goldblatt in analysis of urban enrollments are interesting especially for their strong confirmation of the importance of the "communication" variables: proportions of the population who were in-migrants. Starting from the argument that the greater these proportions the less the urban social-density (to use Durkheim's concept) and the larger the minorities who were not effective participants in modern-oriented information fields, Goldblatt hypothesized that high proportions of in-migrants would be associated with *lower* urban enrollments and school continuation rates. This hypothesis was unambiguously validated in all equations that included either earnings in manufacturing or proportions in white-collar occupations (as proxies for local access to jobs and earnings promising relatively high returns to schooling). The coefficients on the migration variable were quite stable with changes in other variables included.

The other feature of this study that may be of special interest is that in which two-stage regressions are involved. In her initial study, Goldblatt used child-employment as proxy for area differences in foregone earnings or employment opportunities. The sum of enrollment and employment rates for children (aged 8-11) was sufficiently below 1.0 to give

plenty of room for variation in relationships, and the child-employment variable contributed very little in the urban equations though it was of major importance in the rural populations. However, even setting aside statistical bases involved in this use of the child-employment variable, identification problems reappear; the numbers actually employed reflect an adjustment between demands for child labor and supplies of it, supply in turn reflects in part, at least, the schooling decision itself, and that decision again is influenced by demands for child labor. Accordingly, Goldblatt is now exploring a model in which a first-stage equation relates the amount of child-employment to independent variables chosen specifically as relevant to levels of demand for child labor, but with no relevance on the supply side. The child-employment rates predicted by such an equation are then entered in a second stage as the foregone-earnings proxy in the equations for enrollment rates.

### *Components of Productivity Returns to Schooling*

Despite a considerable accumulation of economic studies of research and development and the efforts of econometricians (as yet less than successful) to at least identify "technical change" as something other than a residual in the analysis of economic growth, work directed specifically to the identification of effects of schooling on capacities for or propensities to adaptive and innovative behavior in economic life has only begun.

Some of the most interesting work is thus far largely in the mathematical-theoretical vein. For example, in his exploratory theoretical work on relationships between learning functions and earnings streams, Ben-Porath (1967, 1968) posited and partially tested a "neutrality" hypothesis with respect to relationships between learning-to-learn and learning-to-earn with successive human investments over the life cycle. Nelson (1964) and Nelson and Phelps (1966) have pressed the thesis (argued persistently by sociologists at the Comparative Education Center at Chicago, as well as in my treatments of education in decision theory and in economic history) that the major contribution of schooling to economic growth is in the potential of the better schooled men for adaptive behavior, and the consequent support of innovative change that arises when a large fraction of the population is well-schooled. This inference can be drawn directly also (by quite different but complementary argument) from applications of economic theories of decision making under uncertainty to human-investment decisions.

Whereas Nelson, thinking of employees in larger enterprises, focused on their capacity for adjustment and adaptation to new production equipment and new processes introduced by management, Welch (in press) is concerned with effects of education on what he terms *allocative ability*, which is most important at managerial levels or among independent entre-

preneurs. Allocative ability is seen explicitly in a dynamic context. In empirical application to agriculture, it refers to the capacity of the farmer continuously to alter his output and input combinations in pace with changes at the frontiers of agricultural technology. Empirically, allocative ability in this sense incorporates both a communication or information-field element and an adaptive-innovative aspect of behavior. (This point was argued by Bruce Harker, 1969, in a preliminary paper reporting on the design of his dissertation relating to schooling, communications, and innovative behavior among Japanese farmers.) However, Welch did not go into this distinction, though he makes use of information variables at what Hågerstrand would call the "public" (in contrast to private information-field) level.

Welch's empirical analysis for agriculture rests upon a more general mathematical-theoretical model in which he distinguishes the so-called "worker" from the allocative elements in what education contributes to productivity. The "worker" coefficient would be the one identified in a conventional multiple-regression model in which the marginal contribution of education to output is measured by a partial derivative, holding constant both the output mix and all other inputs into the production process.\* If educational also affects the other inputs and/or the uses to which they are put (the output mix), it gives rise to an interdependency problem analogous in some respects to the problem with which Tolley and Olson (1969) were concerned. The marginal product attributable to education will be measured correctly, by Welch's argument, only by taking the *total* derivative, allowing for the effects of education on other inputs. He showed deductively that "when the marginal product of education is treated as a partial derivative, the bundle of 'other things' held constant is crucial." In developing this analysis (which he does with comparatively simple mathematics) Welch distinguished (a) the "worker effect," which is the effect picked up in an "engineering production function" that holds output mix and other inputs constant, (b) the effect via allocation of inputs among competing uses, which would be picked up, along with the "worker effect" in value-added production functions, and (c) the effect via improved selection of inputs, which will be included, along with (a) and (b), only in a *total* derivative model.

Welch selected agriculture (by States) for empirical application of his analysis of "Education in Production." This is advantageous not only because of the relative accessibility of data, but also because agriculture has been a dynamic industry in the United States, and a much larger proportion of those engaged in agriculture have scope for and must continuously make allocative decisions. Also, as Welch pointed out, jobs performed by persons with differing education are much more differentiated in manu-

\* In fact the more usual models hold inputs constant but allow for variations in output mix, since they use gross revenue (or value added) as the dependent variable.



facturing than in farming. These concepts are carried into empirical application in an analysis of the determination of earnings ratios in agriculture (that is, earnings of college relative to high school graduates, of college graduates relative to functional illiterates, with only 1-4 years of schooling, and of high school graduates relative to the functional illiterates). Independent variables include proportions of each educational class among men engaged in agriculture (skill ratios), non-labor inputs, state expenditures on research per farm, days per farm of time of extension personnel, and a control dummy variable for proportions of farmers who are non-white. The empirical analysis includes examination of factor substitutabilities and of the implied correct or incorrect specification of relationships in different forms and combinations of these variables. Without direct reference to this methodologically rich study, educators will be interested in its major findings. Relative wages for persons who have not attended college are determined by skill ratios only. By contrast, for college graduates a large part of the contribution of education to farm earnings must be attributed to ability to adapt readily to the dynamic pressures in agricultural developments—to continuously keep pace with improvements in available inputs and in the allocation of those inputs among various uses.

The levels of aggregation at which studies are conducted can be decisive for estimates of marginal productivity of education, as was brought out by Welch in relation to his three-way classification of components in the contribution of education to farm productivity. That condition is stressed also by Chaudhri (1968) in his important study of returns to education in Indian agriculture. Chaudhri and Welch join in their common stress on the importance of taking account of effects of education on the input mix selected by the farmer, and hence on the distinction between the "worker effect" of schooling and the input-selection part of Welch's allocative effect. Their interpretations of the aggregation effects on educational coefficients are not quite the same, however.

In the course of his analysis Welch pointed out that previous studies concerning the role of education in returns to the human factor in American agriculture had obtained conflicting results because they used different levels of aggregation. One of these studies, which used observations by States, found schooling to be an important source of variation across States in agricultural productivity. Another, using data by countries, found negligible associations between agricultural productivity and the education of the agricultural population. Welch suggested that the observations on States gave more scope for variations in output mixes, and hence a better chance to observe the allocative components in returns to education. (Neither study allowed for effects of education on choices among other inputs.) The greater variation in output mixes across States is evident enough, and in purely statistical terms this would raise the expected coefficients on education.



However, I have some reservations about Welch's interpretation, since I would question how far the interstate variations in output mix are relevant to analysis of the effects of education on farmers' choices of output mix. In short, again there is an identification problem.

Chaudhri's serious study (1968) of returns to education in Indian agriculture concentrates on examining how different levels of aggregation (and different sets of villages or districts) affect findings and interpretations. Since he used gross revenue as his measure of output, as is usual in studies of production functions, Chaudhri picked up the allocative effects of education via output mix; as Welch pointed out, those measures (at whatever level of aggregation) are not purely "worker effects" even with inputs controlled. Chaudhri stressed a theme that Welch ignored—the "spill-over" or "externality" effect that may be picked up at higher levels of aggregation. This spill-over effect is essentially a communications phenomenon; it might be restated in terms of Hägerstrand's construct, to argue that diffusion of schooling progressively enlarges and intensifies the transmission of modernizing messages through informal, face-to-face information fields, with results for productivity of whole farming communities that would never be predicted by looking simply at within-community differentials in productivity of individuals as related to their schooling.

That changing levels of aggregation can substantially alter statistical findings is hardly open to doubt. Interpretations of this phenomenon raise complex issues. I can not predict where this will lead, but it is clear that the issues raised are not going to die quickly. They are too important for that, in both their substantive and methodological implications.

### *Economists as Educational Planners: The State of the Art*

Methods applied by economists at one time or another to educational planning may be classified in four categories:

- 1) *Manpower planning models.* These are basically alike in their focus on projection of "manpower requirements" for some future target date. Those "requirements" are estimated as a function of projected national income (or rates of growth) assuming fixed input-output coefficients (zero elasticities of substitution in production), usually with adjustments for hypothesized increases in productivity per member of the labor force. Typically a first step is estimation of the "required" occupational mix, from which the estimates of educational requirements are then derived. In some cases, however, the procedure is to go directly to educational requirements without considering occupations at all. The results obtained from manpower-planning models are extremely sensitive to the hypothesized increases in productivity

per member of the labor force. Equally important, they incorporate no assessment of costs or returns to alternative lines of public action or policy. These models concentrate directly upon change through time. They are of course entirely, in the first instance, demand-dominated models.

2) *Input-output models of flows through the educational system.* Here the stress is on the supply side; indeed, some models assume that the supply of skills from the educational system is independent of any sort of economic decision-making. This is the origin of the semantic paradox that the private demand for schooling (the demand of individuals, government action quite aside) has come to be designated in some circles as the "social demand." A number of computable models for these input-output analyses of flows through the schools exist. These models are useful for demonstrating where bottlenecks will occur under pressure of popular demand for more schooling or how far constraints in supply may impede realization of target-date manpower goals. (The models are used in combination with manpower-requirements forecasting for just this purpose.) When projected results look impossible or unfeasible, an iteration procedure can be used to match the demand and supply sides of the manpower-development process. But these also are fixed-coefficient models that lack any criteria for comparing the advantages and disadvantages of alternative policies.

3) *Social benefit/cost analyses.* These are essentially rate-of-return or present-value methods applied to analysis of investments in the expansion of one or another level of schooling, type of curriculum, or scheme of compensatory education (for either children or adults.) Whereas the manpower-planning and educational-flow models are "quantitative" only and start with the simplifying assumption of fixed coefficients and concentrate on estimates of changes over time, the benefit/cost models start from opposite orientations. Their central theme is, as the label implies, to compare benefits with costs for one versus another investment in human beings, or of such investments versus other uses of resources. The observations from which benefit/cost estimates are derived refer in the first instance to a particular point in time (using age cross-section data to get a first approximation to cohort expectations). Unless some further adjustment is made, this procedure implies either (a) an assumption of perfect elasticity of substitution among skills in the face of major changes in the skill mix as education expands (the exact opposite of the manpower planner's position) or (b) an assumption that compensating changes in supplies of other resources and in technologies and demands for skills will maintain the initial productivity differentials between categories of skill.

4) *Linear programming models.* These models maximize attainment of some "objective function" subject to linear constraints and with specific discontinuous restrictions upon resource availabilities or, turning the prob-

lem around, they minimize the costs of attaining these results; the one side of the solution is the "dual" of the other. Usually linear programming models specify bounds that constrain solutions in other ways also, and the bounds may include constraints associated with goals other than that specified in the "objective function." While linear programming is essentially static and has found its main applications at a micro-level, in the business enterprise, the principal applications in educational planning have been dynamic and of national scope. In all such applications an "objective function" derived from one or another benefit/cost model has been combined with one or more of the assumptions that normally characterize manpower-requirements and input-output flow models. In some of the (few) examples of linear programming for national planning of education, the ways of thinking of the manpower planner seem to predominate (Adelman, 1966); in others the neo-classical human-investment and benefit/cost modes of thought predominate (Bowles, 1965). But whatever the balance in linear programming models, there is a critical split between the first two and the last two sorts of model in that the manpower planning and the input-output flow models are quantitative only, whereas the benefit-cost and programming models incorporate some sort of optimizing or sub-optimizing analysis. (The standard work in linear programming in economics is Dorfman, Samuelson, and Solow, 1958.)

For particular projects, which constitute only a small part of an educational system, the simpler benefit/cost analyses can tell as much as will more elaborate procedures. However, when concern is with investments on a vast scale, entailing substantial changes in the structure of the educational system and in subsequent skill mixes of the population, it becomes necessary to combine elements of the manpower techniques with the benefit/cost approach—as in the linear programming efforts. Though they are extremely diverse, the better among these make use of tests for sensitivity to assumed elasticity coefficients in combinations of productive factors both within educational systems and in other parts of the economy. An important finding has been the high sensitivity of results to assumptions concerning those elasticities and associated "shadow prices."

The various major sorts of planning models and different assumptions within each can and do yield markedly conflicting results in recommendations for educational policies (Bowles, in press; Psacharopolous, 1968). Ingenious and enlightening as they may be, these endeavors demonstrate that at this level of planning economists are engaging in an art, not in a science. Whether the educator is facing a policy problem on a national or on a micro-scale, the economist's contribution to the decision will be maximized not by handing out "recipes" or "answers" but by sitting down with the educator to work through an analytical decision process in which there is continuous feed-back between respective specialties.



## Bibliography

- Adelman, Irma. A Linear Programming Model of Educational Planning :A Case Study of Argentina. *The Theory and Design of Economic Development*. (Edited by Irma Adelman and Erik Thorbecke.) Baltimore, Md.: The Johns Hopkins Press, 1966. Pp. 385-411.
- Becker, Gary S. Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy* (Supplement on Investment in Human Beings) 70, No. 5, Part 2: 9-49; 1962.
- Becker, Gary S. *Human Capital*. New York: Published by National Bureau of Economic Research; distributed by Columbia Univ. Press, 1964.
- Ben-Porath, Yoram. The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy* 75: 352-65; 1967.
- Ben-Porath, Yoram. *The Production of Human Capital Over Time*. Paper presented at the Conference on "Education and Income" of the Conference on Research in Income and Wealth, Madison, Wis., Nov. 15, 16, 1968. Jerusalem: Dept. of Economics, Hebrew Univ. (Mimeo.)
- Boulding, Kenneth E. Economics as a Moral Science. *American Economic Review* 59: 1-12; 1969.
- Bowles, Samuel. *The Efficient Allocation of Resources in Education: A Planning Model with Application to Northern Nigeria*. Doctor's thesis. Cambridge, Mass.: Dept. of Economics, Harvard Univ., 1965. Pp. 412-17. (Mimeo.)
- Bowles, Samuel. *Educational Production Functions*. Paper presented at the Conference on "Education and Income" of the Conference on Research in Income and Wealth, Madison, Wis., Nov. 15, 16, 1968. Cambridge, Mass.: Dept. of Economics, Harvard Univ. (Mimeo.)
- Bowles, Samuel. *Planning Education for Economic Growth*, in press.
- Bowles, Samuel and Levin, Henry. The Determinants of Scholastic Achievement. *Journal of Human Resources* 3: 3-24; 1968. (See also continuations of this discussion by several participants, pp. 384-400.)
- Bowman, Mary Jean. Schultz, Denison and the Contribution of 'Eds' to National Income Growth. *Journal of Political Economy* 72: 450-64; 1964.
- Bowman, Mary Jean. The Costing of Human Resource Development. *The Economics of Education*. (Edited by E. A. G. Robinson and J. E. Vaizey.) New York: St. Martin's Press, 1966.
- Bowman, Mary Jean and Anderson, C. Arnold. Concerning the Role of Education in Development. *Old Societies and New States*. (Edited by Clifford Geertz.) Glencoe, N.Y.: Free Press, 1963. Pp. 247-79.
- Chaudhri, D. P. *Education and Agricultural Productivity in India*. Doctor's thesis. Delhi: Delhi School of Economics, Univ. of Delhi, 1968. (Mimeo.)
- Coleman, James S. et al. *Equality of Educational Opportunity*. Washington, D.C.: U.S. Office of Education, 1966.
- Denison, Edward F. *The Sources of Economic Growth in the United States and the Alternatives Before Us*. Supplementary Paper No. 13. New York: Committee for Economic Development, 1962.
- Denison, Edward F. *Why Growth Rates Differ*. Washington, D.C.: The Brookings Institution, 1967.
- DeWitt, Nicholas. *Costs and Returns in Education in the U.S.S.R.* Doctor's thesis. Cambridge, Mass.: Dept. of Economics, Harvard Univ., 1962.
- Dorfman, R.; Samuelson, P. A.; and Solow, R. M. *Linear Programming in Economic Analysis*. New York: McGraw-Hill, 1958.
- Feldstein, M. S. and Fleming, J. S. *The Problem of Time Stream Evaluation; Present Value versus Internal Rate of Return Rules*. Bulletin No. 26. Oxford: Oxford Univ. Institute of Economics and Statistics, 1964.
- Frisch, Ragnar. *Statistical Confluence-Analysis by Means of Complete Regression Systems*. Oslo: Univ. Institute of Economics, 1934.
- Goldblatt, Phyllis K. *Education in Relation to Social and Economic Change in Mexico*. Doctor's thesis. Chicago: Dept. of Education, Univ. of Chicago, 1968. (Offset.)



- Griliches, Zvi. Notes on the Measurement of Price and Quality Changes. *Models of Income Determination*. National Bureau of Economic Research, Studies in Income and Wealth, Volume 28. Princeton, N.J.: Princeton Univ. Press, 1964.
- Griliches, Zvi. Notes on the Role of Education in Production Functions and Growth Accounting. Paper presented at the Conference on "Education and Income" of the Conference on Research in Income and Wealth, Madison, Wis., Nov. 15, 16, 1968. Cambridge, Mass.: Dept. of Economics, Harvard Univ. (Offset.)
- Hagerstrand, Torsten. *Innovation Diffusion as a Spatial Process*. (Translated from Swedish by Allan Pred.) Chicago: Univ. of Chicago Press, 1968.
- Harker, Bruce. *Education and Communication in the Modernization of Rural Japan*. Paper presented at a workshop in Comparative Education. Chicago: Comparative Education Center, Univ. of Chicago, 1969. (Ditto.)
- Harvey, Valerien. *Economic Aspects of Teachers' Salaries*. Doctor's Thesis. Chicago: Dept. of Education, Univ. of Chicago, 1967. (Offset.)
- Hirschleifer, J. On the Theory of Optimal Investment Decision. *Journal of Political Economy* 66: 329-52; 1958.
- Hollister, R. *A Technical Evaluation of the First Stage of the Mediterranean Regional Project*. Paris: Organization for Economic Cooperation and Development, 1966.
- Hood, W. C. and Koopmans, T. C. (editors). *Studies in Econometric Method*. New York: Wiley, 1953.
- Jorgenson, Dale W. and Griliches, Zvi. The Explanation of Productivity Change. *Review of Economic Studies* 34: 249-83; 1967.
- Kendall, M. G. and Stuart, A. *The Advanced Theory of Statistics*. Vol. 2. New York: Hafner, 1963.
- Mincer, Jacob. On the Job Training: Costs, Returns, and Some Implications. *Journal of Political Economy* (Supplement on Investment in Human Beings) 70, No. 5, Part 2: 50-79; 1962.
- Nelson, Richard R. Aggregate Production Functions. *American Economic Review* 54: 575-60; 1964.
- Nelson, R. R. and Phelps, E. S. Investment in Humans, Technological Diffusion, and Economic Growth. *American Economic Review Proceedings* 56, No. 2: 69-75; 1966.
- Organization for Economic Cooperation and Development. *Statistics of the Occupational and Educational Structure of the Labour Force in 53 Countries*. Paris: the Organization, 1969.
- Psacharopoulos, George. *An Economic Analysis of Labor Skill Requirements in Greece*. Doctor's thesis. Chicago: Dept. of Economics, Univ. of Chicago, 1968.
- Schultz, T. W. Capital Formation by Education. *Journal of Political Economy* 68: 571-83; 1960.
- Schultz, T. W. Education and Economic Growth. *Social Forces Influencing American Education*. Chicago: National Society for the Study of Education, 1961.
- Tolley, G. S. and Olson, E. *The Interdependence between Income and Education*. Chicago: Univ. of Chicago, 1969. (Offset.)
- Weiss, Randall D. *The Effect of Education on the Earnings of Blacks and Whites*. Discussion Paper No. 44. Cambridge, Mass.: Program on Regional and Urban Economics, Harvard Univ., Apr. 1969. (Mimeo.)
- Welch, Finis. Education in Production. *Journal of Political Economy*, in press.

## AUTHOR

MARY JEAN BOWMAN Address: University of Chicago Title: Professor of Economics and Education Degrees: B.A., Vassar; Ph.D., Harvard Univ. Specialization: Economics of Education.

## 8: SOCIOLOGY OF EDUCATION

JOHN HARP\*

STEPHEN RICHER

*Carleton University*

Any discussion of the methods employed by sociologists in the study of education must begin with a recognition of the perspective being used. Our discussion must touch upon both the conceptual basis for the subject matter of sociology and the salient character of the methodology used in its study. The juxtaposition of these traditional areas of theory and method within sociology is more than an academic exercise. It provides a framework for a more systematic and comprehensive review of the methods of research used within the sociology of education. Contrary to popular belief there is no one sociological perspective, but rather a range of theoretical approaches are used to study social phenomena. Smelser (1969, p. 5) supported this position when he stated that "sociology, by comparison with some other sciences, lacks a single accepted conceptual framework."

### *Perspectives and Research Methods*

It is now generally accepted that for any given research problem the methods and techniques employed are closely related to the investigator's originating question and perspective. Methodologists of various persuasions have commented on the mutual relationship between theory and fact or theory and research. For example, Nelson Goodman (1964, p. 10) wrote:

Most of our theories are neat approximations, less controlled than inspired by evidence. Failure to transcend the detailed results of observations and experiment is as bad as playing too fast and loose with them. Theory and fact have to be adjusted to each other. The able scientist develops a keen sense of when to yield to recalcitrant observations by modifying his theory and when to blame them on dirty test tubes and faulty instruments.

We subscribe to the established notion of a complementary relationship between theory and research. *Theory* is used here to refer to statements of relationships designed to offer rational explanations about some domain

\* The authors gratefully acknowledge the constructive criticism of the participants in a graduate seminar on educational research held at Cornell University in the summer of 1969.

of interest. The choice of a particular domain by the investigator serves to designate the class of objects to which the theory refers, such as social interaction, collectivities, institutions, etc. A separate but related consideration concerns the unit of analysis, about which more will be said.

### *Research Methods as Social Activity*

Theory in the social sciences receives primary emphasis as the basis for constructing general explanations. This is the *raison d'être* of all scientific activity. However, one should not overlook the fact that theory testing is a social enterprise. The implications of treating methodological rules as social norms bear closer examination. First, there are certain guiding ideals of contemporary empirical science or certain agreed upon general rules of acceptability. These rules provide a basis for determining what statements are to be considered scientific. Although there is a lack of agreement among social scientists about something called "the Scientific Method," discussions of a shared scientific approach are received with greater unanimity. Most sociologists, we believe, subscribe to a catholicity of methods apropos the situation, subject matter, etc., while recognizing their relationship to more general and inclusive criteria. The criteria comprise basic rules of acceptability for scientific statements. These rules serve principally an eliminative function and since they are general rules, they must always be translated by approximation into norms of practice. In this respect a scientist's rules of acceptability are not unlike other ideals of social action. The efficacy of these norms can be judged by the number of *ad hoc* rules required for testing single propositions. Research conducted in compliance with these rules becomes the systematic, controlled and empirical investigation of hypothetical propositions dealing with the presumed relations among social phenomena (Sjoberg, 1965, p. 10).

This sociology of science approach to research methods, although necessarily sketchy and incomplete in its present form, does suggest some questions for a sociology of education. First, one can look at the variety of research within the field and offer explanations for the concentration of research interest on particular problems to the neglect of others. In addition, viewing research methods as norms should enhance one's understanding of certain contradictions in specific research procedures. As constructive critics, researchers are interested in knowing why deviations from norms exist as well as "how informal research procedures are brought to the level of conscious reflection" (Sjoberg, 1965, p. 28).

In the substantive areas of educational research, several recent reviews of the literature show the interest afforded higher education and the problem of class, ethnic and racial differences. Clark (1965) observed that:

The growth of mass higher education, the concern for the relation

of education to national power and prosperity, the interest in the reform of secondary education, and the more recent interest in the education of children from socially and culturally deprived homes are among the powerful external forces leading to the rapid growth of interest and work in the sociology of education.

The pervasive character of an ideology of equal opportunity and the continuing role assigned to education as an avenue of mobility in American society could not help but influence the direction and character of educational sociology. Education is invariably viewed in the normative context of a particular society. Durkheim (1956, p. 71) defined education as follows: "For each society education is the means by which it secures in the children the essential conditions of its own existence." These conditions are normative. In a more limited context, education in an industrial society concerns those formal and systematic approaches to social life that were once the exclusive province of family and community. Since social scientists are usually a part of the society, their choice of research problems is never a purely random phenomenon. Halsey (1968, p. 39), however, saw the social scientists' ability to transcend the concerns of social policy for human betterment as a significant factor in the development of research on educational organizations:

This development owes much to the rise of a group of American and European sociologists who have been determined to treat education in the same terms as any other social institution, that is as an object intrinsically worthy of scientific study, apart from any concern for social policy of human betterment.

Perhaps this explains why there is a paucity of research on educational organizations for it appears that both the quantity and character of research on education is influenced by the values of society.

### *Conceptual and Methodological Parameters*

A majority of the sociological research on education reflects the social concerns of certain collectives in society and at the same time reinforces the strong melioristic tradition of American sociology. However, that tradition is interpreted by a sociology unsympathetic to any social determinism. The salient character of American sociology is often defined as *voluntaristic nominalism*. Wolff (1959, p. 580) offered a succinct definition: ". . . the assumption that the structure of all social groups is the consequence of the aggregate of its separate component individuals and that social phenomena ultimately derive from the motivations of these knowing, feeling and willing individuals."

Seeman (1968, p. 805), in his review of the current International Encyclopedia of the Social Sciences, commented upon this: "In my view



this encyclopedia reports on a great transformation, one that is both thorough going and elusive and concerns the society as well as the social sciences. They have both become psychologized."

Although there has been a tendency to focus on the individual and to study informal structure within education, this does not mean that social structure has been ignored (see Ramsoy, 1963). Quite the contrary; in some of the research completed to date, although the individual was recognized as the unit of analysis with individual attitudes or behaviors as dependent variables, social structure was used as an independent variable (e.g., effect of family, school, community, etc., on educational aspirations and achievement). However, these structural variables were frequently measured by aggregating individual characteristics.

The important distinction is not that data are usually obtained from individuals, for it is generally understood that individual attributes can become system properties, but rather that insufficient attention is often given to the isolation and study of non-reducible structural variables. This unique and salient characteristic of American sociology derives from certain meta-sociological assumptions that afford primacy to psychological realism. The latter position was challenged by Williams (1969, p. 27) who offered a viable alternative for sociology:

Social interactions are not defined for sociology by the muscles, nerves, glands involved, nor by the particular intentions and motives of actors but by the distinctive structuring involved, for example the consensual character of communication and mutual or asymmetrical packaging of units of interaction by actors—their social construction of reality.

Social interaction is surely the conceptual core of the discipline for it represents the most inclusive sociological domain. One is reminded of Durkheim's declarative statement: (1956, p. 35) "As a sociologist, it is above all as a sociologist that I shall speak to you of education." The social analyst is subsequently led, like Durkheim (1956, p. 35), to examine "the very nature of the civilization that education transmits and the mechanisms that it employs to transmit it." Social interaction even in this sense is too broad to guide our review of methods. We must specify the various domains of interest used in the sociology of education. Again, borrowing in part from Williams's (1969) discussion, we offer social organizations, collectivities, and institutions as subdomains worthy of sociological analysis.

Institutions are defined as sets of obligatory norms. The specific values or interests provide the foci for these norms and serve to distinguish among traditional institutional areas in society, such as family, education, and economy. There is a strong normative component inherent in this defini-

tion; this component adds a degree of continuity to the study of social interaction. In contrast, social organizations are primarily patterns of social relationships containing a hierarchial ordering of roles directed toward goals. Finally, collectivities are aggregations of statuses or roles, but lack the pattern of social relations that characterizes social organizations.

The choice of a particular domain of interest by an investigator is usually preceded by an originating question that will serve to clarify what he is trying to do and how he proposes to proceed. All research begins with the investigator's originating question which limits his analysis and greatly influences the methods used. Blau (1969, p. 45) clarified this point by specifying the two general objectives that have directed sociological inquiry: "... the objectives of sociology include not only the explanation of human behavior in terms of social conditions that affect it but also the explanation of why these conditions rather than others came to characterize the social structure."

These objectives have direct implications for the choice of a unit of analysis. Studies designed to determine the attributes of social structure and their interdependence use collectivities rather than individuals as units of analysis (Blau, 1969). The point to be made here, although patently obvious to some, is not often acknowledged in discussions of methodology by sociologists. Methodological approaches vary with the nature of the originating question. If one takes social conditions as given and examines how they affect individuals, this is a markedly different type of exercise than one in which investigators try to explain how these structures developed in the first place.

Given certain conceptual interests for sociology, the dominant character of the methodology used in their exploration remains to be described. In other words, what is the nature of the methods used by sociologists to study institutions, organizations and collectivities? We must establish for our methodology definite limits to what Smelser (1969, p. 35) described as "an attitude of permissiveness for a variety of theoretical and empirical activities." At the most general level limits are specified by certain rules of acceptability. When research is viewed as controlled observation in the search for general explanatory theories, research approaches can be classified as *descriptive-analytical* or as concerned with *testing hypotheses* that have been theoretically deduced.

The former category suggests the use of data and quantitative techniques, thereby eliminating purely discursive accounts of social phenomena. The empirical character of modern sociology is partly based on the following logic (Grebenik and Moser, 1962): "When you can measure something you are speaking about and express it in numbers you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory

kind." Studies designed to test hypotheses are necessarily concerned with controlling for extraneous variance and sources of invalidity. Research of this kind can vary widely in design but the majority of sociological research is non-experimental.

Having specified several domains of interest and two rather elementary but basic research approaches used in sociological research, we are ready to explore their ability in classifying sociological research on education (see Table 1). Needless to say, complete agreement on the units we have chosen does not exist at this time in the discipline.

TABLE 1  
A CLASSIFICATION OF RESEARCH STUDIES  
IN THE SOCIOLOGY OF EDUCATION

<i>Domains of Interest</i>	<i>Research Approaches</i>	
	<i>Descriptive-Analytical</i>	<i>Hypotheses Testing</i>
Social Organizations	Social structure and function of the school	Comparative analyses of educational organizations
Institutions	Educational institution in a societal context	Education and society: Comparative institutional analyses
Collectivities	Descriptive analyses of social climates and student subcultures	Sociology of Learning

#### *Units of Analysis*

A methodological problem closely related to the dimensions used in the paradigm involves the unit of analysis. Since research initiated within any of the three domains of interest may vary with respect to the unit of analysis, this can have a profound influence on methodological approaches. The investigator can choose as his unit of analysis either individual or social units, each having some specific referential system (e.g., secondary school students, urban high schools). Methodological problems arise from two major areas: (1) the assumption that relations between variables at one level hold at another, sometimes referred to as the *displacement of scope problem* and (2) the mixing of levels within the same hypothesis, e.g., collective variables predicting specific kinds of individual behavior. An example of the former problem is the attempt to use generalizations about interpersonal relations as the theoretical basis for the research on inter-

organizational relations. Some clarification of the kinds of measures in actual use is required to identify these problems in greater detail.

The classification of the kinds of properties of collectives as defined by Lazarsfeld and Menzel (1961, p. 422) is helpful in this regard: (a) analytical—a summary measure of the characteristics of individuals, (b) structural—summary measure of the relationship among members, and (c) global—properties not based on information about individual members. Examples of each of these kinds of measure are found in research on social climates. Structural measures are often used in analysis of social interaction networks, and global measures can be used in classifying school systems, for example public versus private. A majority of the sociological research on education involves the use of analytical and structural variables.

The individual is often the unit of analysis in sociological research and individual social behavior, the dependent variable. Much of the research in the sociology of learning can be so characterized. Indeed Blau (1969, p. 53) suggested that this is typical of most empirical research in sociology—the methodology of sociology of education follows the modal pattern for sociology, namely the tendency to treat “the conditions of social structure as given and examine their consequences for individual behavior”. We discuss later the intriguing problem of structural or compositional effects in this area. Research on education has provided investigators with both data and content that permit isolating the structural and individual effects of variables.

In summary, once a researcher has laid claim to a particular domain of interest, research can be viewed as a series of decisions which are all related and which commence with an originating question. The researcher then identifies the major conceptual system to be employed in answering his question. Early in the process an over-all research approach or general strategy is adopted. The use of specific methods or techniques is related to decisions the investigator has made at some earlier period.

The schema presented in Table 1 was designed to depict the major topical areas of research within the sociology of education. In the discussion to follow, however, we emphasize methods employed in studies concerned with hypothesis testing. Essentially those cells in the paradigm representing studies of education as social organizations, the sociology of learning, and studies of education as an institution within some societal context receive major attention. This choice was based on our perception of changes in the direction of research. It is not our purpose to engage in a comprehensive review of the literature in these areas of research, but rather to identify the major methodological problems and where possible suggest what appear to be productive approaches to their solution. We shall attempt in each case to answer two important questions: “What are investigators trying to do in this area of research?” and “How are they attempting to do it?”



We survey some of the major methodological approaches used in representative types of sociological research on education and give recognition to the unit of analysis employed within each type. In the descriptive studies of social organizations, or so-called case studies, the unit is by definition the organization; the investigators attempted to specify organizational properties and less frequently their interrelationships. Systematic comparisons of educational organizations were few. We are not only concerned with a paucity of comparative studies but a deficiency of organizational studies. Reiss (1965) commented on this problem: "Recent attempts to take account of the organizational aspects of the schools are not often germane since they focus either on social structure of the classroom or on some aspect of the social composition of the schools such as its social class composition." This is especially true in the intrasocietal organizational comparisons. That one finds a paucity of studies in this area is sometimes explained by the inappropriate character of survey methods for interorganizational analyses. A burgeoning area includes research on school and society—especially studies employing institutions as a domain of interest. These institutional studies reflect the general trend toward intersocietal research in sociology. Finally we discuss the most prolific research topic within the sociology of education, the sociology of learning.

### *The School as Social Organization*

#### *Case Studies and Comparative Organizational Analyses*

Gross (1959) described the sociology of education as a relatively unexplored but potentially rich area for sociological inquiry. A decade later the volume of sociological research on education causes one to question the current validity of Gross's statement. An important lacuna in this otherwise prolific area of research is organizational studies. When sociologists do use an organizational perspective, social units other than schools are likely to be chosen. Trow (1968, p. 142) stated that "much more has been learned by sociologists about work and industry, politics, the family and religion than about education and institutions." The originating question here is simply: What are schools like as organizations? Studies in which the school was viewed as a social organization are for the most part case studies.

The U.S. system of highly decentralized mass education should facilitate local structural variations ideally suited for comparative analysis (Clark, 1968; Ramsoy, 1963). Nevertheless, case studies have contributed to some understanding of the structural properties of the school as an organization.

In proceeding from originating to specifying types of questions, sociologists have used classical models of bureaucracy to guide their research. There appears to be a trend toward more inductive approaches to describing

organizational properties, in which the properties of the organization are inferred from observations of similar instances in one or more organizations. There is obviously a need for additional research in this area as indicated by Bidwell's (1965) evaluation of current work. "There is no existing study of the prevalence or incidence either of bureaucratic structures or processes in school systems or of their consequences for school system operations. Nor has there been any adequate work on the interplay of bureaucratization and professionalism." Intraorganizational analyses concentrated on the division of labor, the professionalization of school staffs, and the requirements for rationalizing procedures within the system (i.e., hierarchical ordering of roles). Interorganizational analyses often employed more inductive approaches, treating organizations as relatively open systems engaged in exchanges with their respective environments. Some clarification of conceptual frameworks has been achieved, for example the work on mechanisms of authority and control as reflected in the administrative offices of the school. Despite the limited amount of empirical research on this topic, some clarification is evident, for example Clark's (1968) discussion of the bases of authority (public trust, bureaucratic office, collegiability). Litwak and Meyer (1965) offered a theoretical perspective that should prove useful in analyses of school-community relations. What appears to be lacking at the moment are empirical tests of these seminal ideas. On a more abstract level, investigators such as Gross, Mason and MacEachern (1958) used concepts such as universalism and particularism to analyze role expectations and conflicts within the school system. The move from analyses of conflicting processes of bureaucratization and professionalization within the school system to research on role conflicts or "the dilemmas of the teacher's office" appears to be a logical transition. These two research problems are illustrative of Blau's distinction between "(1) the causal interrelations among variables that refer to attributes of the collectivity (social organization) itself, not to attributes of its individual members, and (2) the social interrelations among differentiated positions in the collectivity, such as division of labour or hierarchy of authority."

These are only a few salient theoretical problems involved in the study of schools as social organizations. They serve to specify questions about the structural characteristics of organizations and their interrelationships both within the organization and between the organization and its environment.

Although there is some evidence of cumulative work in this area, total reliance on traditional approaches often encourages a wide range of disparate studies. Investigators tend to use observational methods in case studies and to choose individuals as the unit of analysis when doing surveys. We suggest that survey techniques be used with the organization as the unit of analysis. The insights gained from direct observation, use of informants, analysis of school documents, plus various means of collecting phenomenal data must

be integrated with survey methods if comprehensive research on schools as social organizations is to follow.

### *Observational Versus Survey Approaches*

Discussions of observational versus survey methods abound in the literature. The observational approach is often described as field methods having reference to the observation of people *in situ* (Hughes, 1960) or "all the kinds of techniques which have been employed to examine behavior in naturally occurring groups" (Scott, 1965, p. 262). In descriptive, analytic studies of schools as social organizations, the investigator is likely to collect as much information as possible in order to gain some understanding of the school as an ongoing system. He can use a variety of methods in collecting his data but he is more likely to use a single approach.

The problems that arise in using only one method are particularly illustrated by survey studies. Although surveys invariably employ the individual as a unit of analysis and are primarily concerned with determining the effect of social structure on individual attitudes or behaviors, we shall make brief reference to certain salient characteristics of this approach for purposes of comparison. A more detailed discussion of one or two methodological problems characteristic of survey methods is presented in the section on the sociology of learning.

Typically, a researcher working with survey data never talks to or observes the people whose behavior he is attempting to explain. He statistically *manipulates variables*, documents relationships, and then attempts to interpret the correlations obtained by some kind of intuitive speculation. Suppose he finds a positive relationship between social class and plans to attend a university, a well replicated finding. The next step is to attempt an "interpretation" of the relationship by imagining what there could be about being middle class which would elevate aspirations. Having decided that parental pressure is a salient factor and assuming he has such a measure in the questionnaire, he conducts a partial correlation analysis or uses some analogous technique in an attempt to reduce the original relationship to zero. If this happens the researcher may justifiably claim to have interpreted the social class effect. Unfortunately, such an occurrence is rare in social science research, so that most of the relationship between social class and aspirations will remain intact. At this stage he might introduce several other possible intervening variables with the probable result that none of them, either singly or in combination with others, accounts for the relationship. Here he either blames his measurement procedure or his lack of appropriate test variables and suggests that further research is necessary to document the processes involved.

In short, the researcher is attempting to argue causality *at a distance*. Although survey analysis is no doubt excellent for establishing zero-order

relationships among factors, it fails to provide an understanding of the often subtle processes linking the variables, especially if the researcher is concerned with the operation of interpersonal influences. Not all sociologists have these nagging doubts about survey analysis (Trow, 1968). We return to this topic in our review of research on the sociology of learning.

Continuing our review of "one methodology sociologists," (phrase coined by Whyte, 1965) we found studies that relied completely on observation. Observation has many advantages in the study of organizations since the investigator can record actual behavior when it occurs. Furthermore, since organizations are distinguished from collectives or other social units by their common network of relations and high degree of differentiation, many researchers attempt to understand the structure and social processes within the organization through participant observational techniques. Viewing research methods as an integral part (indeed the norms) of a social science proves useful when applied to field methods. The participant-observer for example, whether he is a sustained researcher or "transitory investigator" (Scott, 1965) is unquestionably a part of the system he is studying. The delicate balance between involvement and objectivity remains a vital issue for participant-observational methods and indeed for all of sociology. Furthermore, it is here that one can observe the process whereby informal and sometimes idiosyncratic research procedures become accepted. We found, for example, a general trend away from highly speculative and discursive interpretations toward more systematic types of observation often accompanied by a change to formalized causal hypotheses and designs.

### *The Future of Case Studies*

Given an emphasis on systematic observation and formal design, the argument for descriptive-analytical studies in which the researcher analyzes a single case very intensively must be reconsidered. Furthermore, recognizing the complexity and diversity of school systems we realize that no one case can stand for all, but case studies often provide insights about processes within the organizations. These can be invaluable in formulating hypotheses for testing in studies using more sophisticated research designs. Given the complex nature of social phenomena and the necessity of choosing a limited number of variables, the judicious use of case studies and observational techniques can play an important role in the early stages of any project. The researcher must be on guard lest his interest in "preserving the unitary character of the social object being studied" (Goode and Hatt, 1952, p. 130) jeopardizes his ability to generalize from his materials. We refer here to the conceptualization of general categories with universal application.

In addition to its utility as an unexploratory device, the case study approach when combined with observational techniques is useful in the study of change (although it is seldom so employed). Investigators studying



organizational change have been quick to point out distinctions between formal and informal structures within organizations. Presumably the informal patterns of interaction within the organization can be extremely relevant in identifying directions of change. Furthermore, the distinction between ideal and real normative patterns is a very old sociological problem which can be clarified through observational techniques.

Although sociologists are still grappling with the uses of time in sociological analysis, the participant observer will probably be acutely aware of the changes occurring within the social unit he is studying. Unfortunately, in a large and complex organization the luxury of sustained participant-observation is simply not feasible. Current interest in sampling of time periods and interaction systems within organizations may provide solutions to this problem.

Suffice it to say that what appears to be needed is not the elimination of case studies *per se* as some have suggested, but a more effective utilization of case studies in conjunction with more sophisticated designs. One could wish for the completion of a series of case studies covering a range of educational organizations, all of which were classified *a priori* on certain relevant structural characteristics. It is with this ideal in mind that we examine certain comparative approaches either intra-or inter-societal in the study of educational systems.

## *Education and Society*

### *Institutional Analyses*

Research on education is frequently part of general holistic approaches to the study of total societies. The use of more holistic approaches has certainly been characteristic of British sociology, as suggested by Floud (1962, p. 225): "It is not surprising that British sociologists should be interested in the effects of the educational system on these movements in the wider society, and that the relation of education to the class structure and of both to the national economy should be the most systematically and fully cultivated part of education in Britain."

The interest in education as an institution within the total society resulted in many descriptive and some analytical studies. One major focus of such studies is on institutional interrelationships, more specifically between education and the economy or education and the stratification system. Current interest centers on the role of education in theories of social and economic development since it is generally conceded that industrial societies require higher levels of formal education. The development and use of modern complex technology places a premium on the continued development of formal educational systems to supply a skilled labor force.

Studies that treat education as an institution and purposely compare societies classified on some important structural variable are rare. One is more likely to find a collection of studies chosen to illustrate the proposition that the educational system of a society is shaped and directed by one or more of the four basic social institutions: family, church, state and economy (Havighurst, 1968). Each of the studies can stand alone and are often, though not exclusively, descriptive in character. A collection of intrasocietal studies places the burden on the reader to make the comparisons among societies. All of these are a far cry from the criteria established by Marsh (1967) for comparative analyses, namely a common design, a relatively large number of societies, and explicit comparisons for this subclass of studies. This ideal is seldom achieved in practice.

*Education and Societal Development.* Cross-national studies are not totally absent from the field of education. Comparative education is hardly a new concept, but it is now defined in the light of both contemporary problems and methods of the social sciences. Brickman (1968) comments on the development of the concept:

The term 'comparative education' is of recent origin and as far as can be ascertained dates from the beginning of the nineteenth century. However, the notion of equating educational methods or indulging in comparative commentaries has been with pedagogy ever since scholars and teachers first came into conflict with ideas different from their own. The interchange of ideas and the dialectics of education have been occurring in a continuing basis ever since scholarly disputations were first recorded.

At the Comparative Education Center in Chicago (Anderson, 1965), comparative education is defined as "the cross cultural study of a) the relationships between education and other aspects of society and b) of the interrelationships among the various aspects of education." Much of the interest in cross-national studies focuses on the problems of societal development. Studies in development education, although relatively new, have experienced many methodological problems. Adams and Farrell (1966) commented objectively about the state of research in development education:

Studies in development education are relatively new, faddish and often weak in theoretical foundations. These studies frequently are interdisciplinary which is to their advantage. They do, however, suffer from both the underdeveloped state of the discipline of education, with its lack of theoretical bases and absence of well developed methodology and from shortcomings particularly in terms of measurement, inherent in the social sciences.

The originating question for this line of inquiry is: What are the ways in which education may be affected by and, in turn, support certain crucial aspects of the development process?

*Problems of Data Collection.* Measurement problems originating with the lack of strictly comparable data from different countries is an Achilles Heel for research of this kind. Published statistics that usually provide a data source are of mixed quality. As one observer described it, the common practice is to admit the data are bad and then proceed as though they were good. Deviations from strict comparability are typical in comparative studies of occupations. The definition of a physician in one country might be equivalent to that of a third-year medical student in another.

When collecting new data for research the problems of cross-national studies are legion. It is now generally recognized, for example, that many of the sources of error are country specific. Once again, this suggests a sociology of science perspective for research methodology. A point made by Coleman (1969, p. 98) in his review of the methods of sociology:

In such research, stories abound with the description of these country specific errors. In one study, it was said the key punching errors reflected differences in national character: Germany showed a total absence of random errors in key punching; whenever an error occurred it was a systematic error that appeared in all cards; in Italy the errors were scattered with apparent Latin abandon.

A cross-national study of school achievement illustrates this problem; the report contains a note stating that the results given for Finland are incorrect because of poor communication (see Husén, 1967).

The seriousness of the error depends upon the purposes for which the data were gathered. For example, the question of over or under estimates can be less crucial when rank ordering countries. It should be recognized, however, that where total populations are involved, and this is often the case in cross-national studies, tests of significance are largely irrelevant. In short, the statistic employed should be appropriate to the level of measurement and accuracy of the data.

A major problem of international surveys is question wording as it relates to conceptual and operational equivalence. The procedure for translation of questions is now fairly standard with the original instrument first translated in the local language and then translated back into the original (Mitchell, 1968). Investigators in this area of research are fortunate to have a relatively wide array of quantitative data from which to choose their indicators, and this largely is due to the efforts of the United Nations. The more popular measures constructed by investi-

gators are: literacy rates; primary and secondary teachers per 1,000 population; enrollment ratios such as primary school as percentage of 5-14 age group; secondary school as percentage of 15-19 age group; primary, secondary and post primary as a percentage of the total population. Two salient problems arise with enrollment ratios—categories in which data are reported can be country specific, and the quality and content of education is usually ignored.

In the final analysis, the validity of a measure and its utility can only be evaluated in terms of the concept one is trying to measure. Many cross-national studies seem to be primarily concerned with correlating a large array of indicators without the benefit of hypotheses. One has the distinct impression that for many analyses, objectivity is to be achieved by extracting theory directly from data. However, there are some encouraging signs of greater conceptual sophistication.

Some researchers who viewed education as a process developed measures of educational inputs and outputs. We have already referred to certain traditional input measures such as number of pupils and teachers. Output measures include achievement test scores and entrance examination scores for admission to higher schooling. Quite recently new measures have been suggested such as the pupil-hour measure (inputs) and the pass-years measure (outputs). The pupil-hours measure is attributed to Adams and Farrell (1966). Several criteria have been suggested for evaluating measures of educational output. For example, it is suggested that the indicators should reflect what is learned and what is taught; a homogeneous concept is sought, preferably one that is comparable within countries as well as internationally. Researchers want answers to the following questions about societies: "How many pupils of what types are selected for education? What behavioral capabilities are achieved, to what degree and by how many?" New measures of educational inputs and outputs are needed in the analysis of education and development. Some studies, for example those in manpower and educational planning, represent what Coleman (1969) referred to as a new development for sociological methods. This new development is the marriage of survey research and demography through an exchange of methods, a trend facilitated by computer programs. Coleman (1969, p. 89) elaborates on the characteristics of this research: "With increasing frequency, such research is going to be carried out, research that combines the educational depth with the scope, representativeness and quantitative final measures that have characterized the best demographic research."

A specific example of the trend referred to above is a new use for scalogram analysis. This technique was originally developed by Guttman for attitude research with the individual as the unit of analysis, but it has recently been extended to measure structural properties of collectives



such as communities and societies. Guttman scales of educational complexity have been developed and tested on a number of societies. Ruth Young (1966, p. 101) cited some important characteristics of scaling when the data are social objects:

The situation in Social Object Scaling, however, is clearly different (i.e., from attitudinal scaling). The items of analysis are not answers to carefully phrased questions but are instead societal traits and institutions. It is a commonplace to say that social institutions are resultant of many interacting forces of many underlying variables. Further, a given institution frequently performs diverse functions. Thus it is quite possible for educational planning institutions to reflect the institutional differentiations of both the educational systems and the planning systems.

This particular tool can provide comparative insights into national development and contribute to theories of social change.

In summary, the comparative study of educational systems across a range of societies is a most vital area of contemporary sociological research. The potential contributions of such studies to the general knowledge of sociology as well as to the important implications of research for programs of change are great. Improvements in the range and quality of data derived from the skillful blending of survey and demographic techniques are likely results of wider use of the comparative approach.

### *Sociology of Learning*

A large volume of American research on education concerns the problem of explaining educational aspirations and achievement. This type of research has typically related some aspect of the student's environment to some measure of his performance. The independent variable is usually some social condition, and the dependent variable represents individual attitudinal or behavioral variables. Research of this kind posits the originating question, "How are certain kinds of social behavior affected by varying social conditions?" Surveys on the influence of such characteristics focused on two major social units, family and school. By strict definition, research of this kind is social psychology. However, it belongs in this review because so much of sociology is precisely of this character and many interesting methodological problems arise from this mixing of conceptual levels.

#### *The Family as Independent Variable*

Sociologists have long been interested in the influence of the family as a socializing agency. Within the sociology of education, a student's family of orientation is viewed as performing a status assignment function.

An individual's social class position is usually measured by the socio-economic status of his family. The higher the family on this factor (as measured by various combinations of parents' education and occupation) the higher the student's aspirations or achievement. In attempting to identify specific sources of influence on the student's educational and vocational aspirations explanatory variables such as family cohesion, presence of role models and social climate of the family have been explored. Also, rival hypotheses concerning parental versus peer group influences on aspirations are subsumed under sources of influence.

Charters (1963) identified a wide variety of aspiration and achievement measures known to be related to socio-economic status (SES). He reported (Charters, 1963, p. 740) that SES predicts:

grades, achievement and intelligence test scores, retentions at grade level, course failures, truancy, suspensions from school, high school dropouts, plans for college attendance, and total amount of formal schooling. It also predicts academic honors and awards, elective school offices, extent of participation in extra-curricular activities and in social affairs of the school, to say nothing of a variety of indicators of success in the informal structure of student society.

Methodologists have been quick to point out that prediction does not constitute explanation. Nor can one variable be viewed as a necessary and sufficient condition for a given social result. Consequently, a large number of intervening variables have been introduced to interpret the results. Boocock (1964) cited four of these intervening variables: values, child-rearing practices, family size, and relationship with teachers. Given the high degree of correlation among these factors and lacking a theoretical perspective, one finds that conflicting hypotheses and *post facto* interpretations abound. Very often the sample size seriously inhibited the use of more than one or two control variables when the analysis was of the multivariate contingency table type. Even when statistical controls were used, the reduction in explained variance was often insufficient to prove conclusive.

There are two recent methodological developments that portend to rescue this area of research from a state of directionless empiricism. They are the result of what Coleman (1969, p. 91) referred to as a new statistical sophistication among sociologists that has brought about "a convergence of statistical analysis based on continuous variables and that based on categorical data." Coleman (1969, p. 91) succinctly described the gains associated with this development, namely that the researcher: "can examine simultaneously the partial relationships of a number of variables to the dependent variable, and thus is far less likely to obtain a fake picture of the true causal structure by failing to control simultane-

ously on a number of relevant variables." To illustrate, the sociologist can now look at the relationship between ordinal variables (e.g., social class) and nominal variables (e.g., drop-out plans) with simultaneous controls on other nominal or continuous factors.

A second methodological breakthrough and one particularly applicable to research on education is the development of causal models through path analysis (Blalock, 1964; Duncan, 1966; Sewell, Haller and Portes, 1969). Those who apply this (relatively new) technique to sociological data argue persuasively that path analysis established the basis for transforming sociological theory and research from qualitative to quantitative. It is of further interest to note that research on education provides both problem and data suitable for path analysis (Coleman, 1969).

### *The School as Independent Variable*

Sociologists have displayed a range of conceptual interests in the effect of school on student behavior. Schools have been characterized by the ideology of their "leading crowds"; the social class composition of their students, their geographical locations—rural-urban; and numerous other dimensions displaying all three measurement properties: (analytical, structural, global). A typical strategy involves using aggregated measures or analytical properties to type schools. The researcher classifies schools as lower or middle class by the status of their students and then makes the necessary logical comparisons, paying particular attention to lower class students in predominantly middle class schools and vice versa. There are two distinct problems associated with using the school as an independent variable, namely the self-selection problem and that of statistical control.

### *Self-Selection Problem*

In non-experimental studies using the school as a unit, there is a methodological difficulty in distinguishing individual from contextual effects (Blau, 1960; Davis, Spaeth and Hudson, 1961). The crux of this problem was summarized by Wilson (1968, p. 83): "However, it remains true that the hypothesis of a contextual effect, where the relevant context is determined by aggregative characteristics of the members, is always vulnerable to the counter-hypothesis of self-selection." That is, what appears to be a school effect might in reality reflect the fact that students attending a particular type of school are different to begin with from those attending a second type. These differences might therefore be the true source of the apparent school effect. The nagging doubt raised by the phenomenon of self-selection is thus always present. Part of the problem rests with the fact that studies which use school level characteristics as independent variables are usually cross-sectional.

In cross-sectional survey research in education, there are typically

two groups of students, one in a school which has a particular characteristic, one which does not. The students are compared on some performance or aspiration measure. The difficulty arises because the two groups have not been randomly assigned to the schools they are attending. The researcher cannot therefore assume the groups were randomly equivalent on all dimensions related to performance before their school experience. One might argue, then, that the school factor is not the "real" reason for any difference in performance, but that students entering a school of a particular type differ initially from those in a second school.

Sociologists have attempted to solve this problem in three ways, none of which is entirely satisfactory. One approach is to statistically partial out the effects of various background characteristics of students. This would help in clarifying whether the relationship between a structural property of the school and student performance was spurious. Variables such as the family factors cited earlier, including SES, educational values, family climate, are likely choices for covariates. Unfortunately, this array of general indicators does not exhaust the range of subtle reasons that individuals can employ in choosing one educational institution rather than another. About all one can say is that those particular variables which are validly and reliably measured and employed in the partialing out process have been controlled. Furthermore, this is true whether one uses partial correlations, multiple regression or analysis of covariance as statistical means of control.

Another strategy is to compare the strength of the relationship in early grades with its strength in later grades. Thus, if there is a greater effect of a school factor for senior than freshmen students, the argument might be made that increased exposure to the school has improved performance. The argument can still be advanced that freshmen are not the same students as seniors. This fact remains despite attempts to limit comparisons to students similar in SES, IQ, etc.

A final strategy involves substituting a longitudinal design for the original cross-sectional design. When combined with a panel study the problem of different individuals being compared is solved but the self-selection difficulty remains. The essential feature is that data are collected at two or more points in time on the same sample (Astin, 1968). The argument is that if one can compare the *change* in performance of the same individuals as exposure to the school increases, this change will reflect a valid school effect.

Two possible solutions to the self-selection problem come to mind. One is to come closer to the classical experimental design. This implies random assignment of students to schools, a procedure which is only possible with total cooperation on the part of educators and school administrators. A less attractive alternative is to include questions on question-



naires which explicitly aim at controlling the selection variable. Asking students whether they chose or were assigned to their school is an example. Comparing students in different institutions who were assigned to their school would eliminate the possibility that they were in that particular institution by virtue of self-selection.

### *Problems of Statistical Control*

A second problem with using school characteristics as independent variables is that of statistical control. This is especially relevant when one is working on school level variables which are aggregated characteristics of individuals, for here the problem of sorting out individual versus school effects is most crucial. The failure to do so may well lead to spurious results (Tannenbaum and Bachman, 1964).

Variously termed "structural analyses," "compositional analyses" and "contextual analyses," the general strategy is to view the individual in a social context derived by combining or aggregating particular characteristics of his fellow students into a school level measure. A prime example is social class or SES composition. The independent variable could be the proportion of "middle class" students in a school. This variable is then related to the individual's aspirations or performance, controlling for the individual's own social class background. Comparisons typically carried out are: (1) lower-class students in predominantly lower-class schools versus lower class students in predominantly middle-class schools and (2) middle-class students in predominantly lower-class schools versus middle-class students in predominantly middle-class schools.

The investigators of the majority of studies in this area report that middle and lower class students perform better in middle-class institutions (Boyle, 1966; Michael, 1961; Wilson, 1959). The specific nature of the causal process involved is said to be a kind of diffusion of achievement values that results from social interaction with individuals holding such values. That is, since middle-class students are known to possess strong pro-achievement attitudes, a student in a school composed mainly of middle-class students is more likely to adopt or continue holding such values than is his counterpart in a lower-class institution. This process was documented by Campbell and Alexander (1965), who found that the relationship between social class composition and college plans disappears when the social class of the student's close friends is partialled out. The causal chain is as follows:

High middle class school → \*Individual's  
friends more likely to be middle class →  
Greater contact with middle class values →  
Higher college aspirations.

\* Arrow denotes a causal relationship in the direction indicated.

Recently, however, these findings have come under attack, notably in the work of McDill et al. (1966). In these studies, no effect of SES composition was shown when the investigators controlled for individual SES and some measure of IQ. These investigators concluded that the school level effect found by Wilson (1959) and others reflects only an unsuccessful attempt to control factors such as family SES and ability.

The conclusion appears reasonable except that contradictory evidence is found in Sewell and Armer's data. Although the authors eliminated the overall effect of SES composition by means of a multiple regression analysis, the effect persisted within certain subgroups in their sample. The latter point received relatively little attention since the authors were primarily interested in the main rather than the interaction effects. The point is, however, that it is precisely the interaction effects which are salient in most recent research on school effects. The significant but controversial Coleman (1966) report on equality of educational opportunity is an example. The main finding shows that among Negro high school students, 80 per cent of the variance in verbal ability is accounted for by family background characteristics, with a corresponding figure of 90 per cent for whites. Within the residual category, school level factors (primarily the composition of the students' background and aspirations) accounted for a greater proportion of the variance among blacks than for the white subpopulations. Similarly, Coleman's study showed that school factors had their greatest effect on those students whose potential occupational and educational success were lowest.

Similar results were reported in a recent study of Canadian high-school students. School level variables had a consistent and pronounced influence on low rather than high IQ students (Richer, 1968). It appears that students less favorably endowed with those social attributes known to be associated with achievement are somewhat more sensitive to environmental factors. School factors may thus serve a substituting kind of function, taking up the slack for those individuals not already highly motivated.

The point to be made then is that researchers in education have tended to deemphasize interaction effect while concentrating on over-all summary effect measures. More detailed, careful investigations of what's going on in various subgroups are needed to answer the question: "Whom does the school affect?" The answer would be fruitful in developing a general theory of school level effects.

### Summary

We have attempted to evaluate the major methodological approaches which typified sociological research in education. Concentrating on several topical research areas, certain problems that arise from the methods currently employed by investigators were discussed:

1. There was a proliferation of case studies of individual educational organizations as opposed to systematic comparative studies of educational organizations. This emphasis has impeded the development of causal hypotheses about variation in organizational structure and delayed inter-organizational analyses.
2. There often appeared to be a relationship between a given substantive area and the method chosen for its study. Studies of schools as social organizations were often case studies in which observational methods were used. Rose (1967) expressed some concern with this point in his discussion of "tautological" proof and its cure. A combination of methods would serve as a more valid test of hypotheses and ensure, to use Rose's (1967, p. 213) terminology, that the method does not "*logically* produce a finding which confirms the hypothesis."
3. There were few studies either intra- or inter-societal, on the relationship between educational institutions and other societal institutions, as recent reviews of the literature revealed. The interplay between school and church, school and economy, school and family, and school and state, especially as these vary among societies, received considerably less than adequate attention. Problems of measurement—reliability and validity—are possible impediments to successful cross-national research.
4. There was a dearth of systematic causal analyses of individual attitudes and behavior. For example, there is little knowledge about the nature of social class effects on educational plans. Developments such as the convergence of categorical and continuous data and the application of path analysis are promising innovations for sociological analysis.
5. Several problems, as yet unresolved, accrued from researchers using the school as an independent variable. Self-selection and statistical control problems are most crucial here. Also, the tendency to deemphasize interaction effects of school level variables is a detriment to theoretical development.

This brief list is in no way exhaustive; it reflects the authors' perceptions of the current state of sociological methodology. The sociology of education continues to provide opportunities for investigators to refine methods in sociological analyses essential for testing theories that can contribute to the solution of practical problems facing educational systems.

### *Bibliography*

- Adams, Don and Farrell, James P. (editors). *Education and Social Development*. Syracuse, N.Y.: Center for Development Education, Syracuse Univ., Mar. 1966.
- Anderson, C. Arnold. *The Program of the Comparative Education Center, University of Chicago. Government Policy and International Education*. (Edited by Stewart Fraser.) New York: John Wiley and Sons, 1965.
- Astin, Alexander W. *The College Environment*. New York: The American Council on Education, 1968.

- Bidwell, Charles E. *The School as Formal Organization. Handbook of Organizations.* (Edited by James G. March.) Chicago: Rand McNally, 1965.
- Blalock, Hubert. *Causal Influences in Non-Experimental Research.* Chapel Hill: Univ. of N.C. Press, 1964.
- Boocock, Sarane S. Toward a Sociology of Learning: A Selective Review of Existing Research. *Sociology of Education* 39: 1-45; 1966.
- Boyle, Richard P. The Effect of High School on Students' Aspirations. *American Journal of Sociology* 71: 628-39; 1966.
- Blau, Peter. Structural Effects. *American Sociological Review* 26: 178-93; 1960.
- Blau, Peter. Objectives of Sociology. *A Design for Sociology: Scope, Objectives and Methods.* (Edited by Robert Bierstedt.) Philadelphia: The American Academy of Political and Social Science, 1969.
- Brickman, William W. March-Antoine Jullien and His Plan. *Jullien's Plan for Comparative Education.* (Edited by Stewart Fraser.) New York: Bureau of Publications, Teachers College, Columbia Univ., 1964.
- Campbell, Ernest Q. and Alexander, C. Norman. Structural Effects and Interpersonal Relations. *American Journal of Sociology* 71: 284-89; 1965.
- Charters, W. W., Jr. The Social Background of Teaching. *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally, 1963.
- Clark, Burton. Sociology of Education. *Handbook of Modern Sociology.* (Edited by Robert E. Faris.) Chicago: Rand McNally, 1965.
- Clark, Burton. Education: The Study of Educational System. *International Encyclopedia of the Social Sciences.* (Edited by David L. Sills.) New York: MacMillan and the Free Press, 1968.
- Coleman, James S. et al. *Equality of Educational Opportunity.* Washington, D.C.: United States Government Printing Office, 1966.
- Coleman, James S. The Methods of Sociology. *A Design for Sociology: Scope, Objectives, Methods.* (Edited by Robert Bierstedt.) Philadelphia: The American Academy of Political and Social Sciences, 1969.
- Davis, James; Spaeth, Joe L.; and Hudson, Carolyn. A Technique for Analyzing the Effects of Group Composition. *American Sociological Review* 26: 215-26; 1961.
- Duncan, Otis Dudley. Path Analysis: Sociological Examples. *American Journal of Sociology* 72: 1-16; 1966.
- Durkheim, Emile. *Education and Sociology.* (Translated by Sherwood D. Fox.) Glencoe, Ill.: The Free Press, 1956.
- Floud, Jean. The Sociology of Education. *Society, Problems and Methods of Study.* (Edited by A. R. Welford et al.) London: Routledge and Kegan Paul, 1962.
- Goode, William J. and Hatt, Paul K. *Methods in Social Research.* New York: McGraw-Hill, 1952.
- Goodman, Nelson. Science and Simplicity. *Forum Lectures: Philosophy of Science Series No. 16.* Washington, D.C.: United States Information Agency, 1964.
- Grebenik, E. and Moser, C. A. Statistical Surveys. *Society, Problems and Methods of Study.* (Edited by A. T. Welford et al.) London: Routledge and Kegan Paul, 1962.
- Gross, Neal. The Sociology of Education. *Sociology Today.* (Edited by Robert Merton et al.) New York: Basic Books, 1959.
- Gross, Neal; Mason, W.; and MacEachern, A. W. *Explorations in Role Analysis: Studies of the School Superintendency Role.* New York: John Wiley and Sons, 1958.
- Halsey, A. H. Education. *International Encyclopedia of the Social Sciences.* (Edited by David L. Sills.) New York: MacMillan and the Free Press, 1968.
- Havighurst, Robert J. (editor). *Comparative Perspectives on Education.* Boston: Little, Brown, 1968.
- Hughes, Everette C. The Place of Field Work in Social Science. *Field Work: An Introduction to the Social Sciences.* (Edited by B. H. Junker.) Chicago: Univ. of Chicago Press, 1960.
- Husén, Torsten (editor). *International Study of Achievement in Mathematics.* Vols. 1 and 2. New York: John Wiley and Sons, 1967.
- Litwak, Eugene and Meyer, Henry J. Administrative Styles and Community Linkage. *Schools in a Changing Society.* (Edited by Albert J. Reiss.) New York: The Free Press, 1965.
- Lazarsfeld, P. F. and Menzel, H. On the Relationship between Individual and Collective



- Properties. *Complex Organizations: A Sociological Reader*. (Edited by A. Etzioni.) New York: Holt, Rinehart and Winston, 1961.
- Marsh, Robert M. *Comparative Sociology*. New York: Harcourt, Brace and World, 1967.
- McDill, E. et al. *Sources of Educational Climates in High Schools*. Washington, D.C.: Dept. of Health, Education and Welfare, 1966.
- Michael, John A. High School Climates and Plans for Entering College. *Public Opinion Quarterly* 25: 585-95; 1961.
- Mitchell, Robert E. Survey Materials Collected in the Developing Countries: Obstacles to Comparisons. *Comparative Research Across Cultures and Nations*. (Edited by S. Rokkan.) Paris: Mouton, 1968.
- Reiss, Albert J., Jr. (editor). *Schools in a Changing Society*. New York: the Free Press, 1965.
- Ramsay, Natalie Rogoff. Public Education in America: A Research Program. *Sociology of Education* 37, No. 1; 1963.
- Richer, Stephen. *Programme Grouping and Educational Plans: A Study of Canadian High School Students*. Doctor's thesis. Baltimore, Md.: Johns Hopkins Univ., 1968.
- Rose, Arnold M. The Relation of Theory and Method. *Sociological Theory: Inquiries and Paradigms*. (Edited by Llewellyn Gross.) New York: Harper and Row, 1967.
- Scott, W. Richard. Field Methods in the Study of Organizations. *Handbook of Organizations*. (Edited by James G. March.) Chicago: Rand McNally, 1965.
- Seeman, Melvin. Review Symposium on International Encyclopedia of the Social Sciences. *American Sociological Review* 33: 805-807; 1968.
- Sewell, William H.; Haller, A.; and Portes, A. The Educational and Early Occupational Attainment Process. *American Sociological Review* 31: 82-92; 1969.
- Sjoberg, Gideon. *Research Methodology: A Sociology of Knowledge Perspective*. A paper read at the Annual Meeting of the American Sociological Assoc., Chicago, Ill., Sept. 1965.
- Sjoberg, Gideon and Nett, Roger. *A Methodology for Social Research*. New York: Harper and Row, 1968.
- Smelser, Neil J. The Optimum Scope of Sociology. *A Design for Sociology: Scope, Objectives and Methods*. (Edited by Robert Bierstedt.) Philadelphia: The American Academy of Political and Social Science, 1969.
- Tannenbaum, Arnold S. and Buchman, Jerald G. Structural versus Individual Effects. *American Journal of Sociology* 69; May 1964.
- Trow, M. The Sociology of Education. *American Sociology*. (Edited by T. Parsons.) New York: Basic Books, 1968.
- Whyte, William Foote. *Toward an Integrated Approach for Research on Organizational Behavior, Industrial and Labor Relations* Reprint Series No. 155. Ithaca, N.Y.: Cornell Univ., 1965.
- Williams, Robin M., Jr. Comments on Smelser's Paper. *A Design for Sociology: Scope, Objectives and Methods*. (Edited by Robert Bierstedt.) Philadelphia: The American Academy of Political and Social Science, 1969.
- Wilson, Alan B. Residential Segregation of Social Classes and Aspirations of High School Boys. *American Sociological Review* 24: 836-45; 1959.
- Wilson, Alan B. Social Class and Equal Educational Opportunity. *Harvard Educational Review* 38, No. 1; 1968.
- Wolff, Kurt H. The Sociology of Knowledge and Sociological Theory. *Symposium on Sociological Theory*. Evanston, Ill.: Row, Peterson, 1959.
- Young, Ruth C. Use of Scalogram Analysis in the Study of Development. *Education and Social Development*. (Edited by Don Adams and James P. Farrell.) Syracuse, N.Y.: Center for Development Education, Syracuse Univ., 1966.

## AUTHORS

- JOHN HARP Address: Carleton University (Canada) Title: Professor of Sociology Age: 41 Degrees: B.S., Univ. of Saskatchewan; M.S. and Ph.D., Iowa State Univ. Specialization: Research Methods; Community Studies.
- STEPHEN RICHER Address: Carleton University (Canada) Title: Assistant Professor of Sociology Age: 28 Degrees: B.A. and M.A., McGill Univ.; Ph.D., The Johns Hopkins Univ. Specialization: Sociology of Education; Ethnic Relations; Child Development.

## 9: EDUCATION'S CHALLENGE TO PSYCHOLOGY: THE PREDICTION OF BEHAVIOR FROM PERSON-ENVIRONMENT INTERACTIONS

JAMES V. MITCHELL, JR.

*University of Rochester*

Past reviews of research methodology in the *Review of Educational Research* have exhibited some definite trends over time. In her review of trends in research methodology for the special RER issue entitled "Twenty-Five Years of Educational Research," Walker (1956, p. 326) described how successive issues on research methodology tended to place more and more emphasis on the "... variety, complexity, and importance of statistical methods for analyzing data gathered by various techniques." It is interesting to conjecture about how the present issue on research methodology will be evaluated some years hence against the background of these past trends and the undoubtedly accelerated trends of the future. This issue departs appreciably from the usual format and requires the author to examine how research pertinent to education is being conducted in his substantive field and to provide a critical analysis and evaluation of a few selected areas of current research endeavor from the point of view of relevance, generalizability, potential, and limitations. The task is one that stimulates a desire to take stock of current conditions and needs in the field and to speculate about desirable directions for the future.

The contemporary social and political context certainly encourages such stock-taking. The challenge that speaks (and sometimes shrieks) of "relevance" can be rejected, ignored, controverted, rationalized, or redefined in less objectionable terms, but it still exists and must be contended with in one way or another. And any socially conscious research worker is certain to have periods of heightened concern and remorse when he considers the enormity of those contemporary social issues in relation to the utility of the armamentarium that psychology can bring to bear upon them. What we need to know about the specific determinants of cognitive deficit in the disadvantaged is staggering, and the educational methods of greatest benefit to its amelioration are still a matter of conjecture and opinion. The turmoil on America's campuses is now being explained *ex post facto* and with vigor by the armchair theorists, but such activity does not hide the psychologists' very evident inability to predict the psychosocial determinants of these events and their ultimate outcomes.

The relevance of the challenge of relevance must itself be evaluated carefully in terms of the present level of development of psychology and the nature of the contribution that can be envisaged, hopefully, for the present or some future time. Such issues will not be debated here. But the political issue of relevance can serve some useful scientific purpose if it serves to remind researchers that scientific or theoretical relevance does not exist unless their work leads to the explanation and prediction of the events with which they are concerned. What is painfully apparent about recent campus unrest is that psychology neither predicted it in the past nor fully understands it in the present. About the best that psychologists can do is to engage in concurrent studies and take note of the fact, for example, that student leaders of the protest movement are above average in scholastic aptitude and parental affluence. Such contributions are less than helpful. The latest in trait inventories or other presumed sources of information about intrapsychic states adds precious little to that understanding and little is expected of them; typically the devotees of the field and their leaders either engage in the aforementioned *ad hoc* theorizing, choose a safer subject to work with, or join with others in a frankly political stance. The net result is neither substantive nor useful.

What these contemporary events seem to suggest most forcefully is that the determinants of behavior need to be sought more often in the characteristics of the environmental context and the interaction of these characteristics with individual traits and abilities, and that a search for individual characteristics *in vacuo* can lead only to partial understanding or no understanding at all. The processes used by the individual to "come to terms with the environment," as Goldstein (1939) wrote, are processes that have much to say about that individual's values, traits, and characteristic modes of functioning. Faced with the almost cataclysmic impact of contemporary social and political events, it should not be difficult to accept the importance of this interaction for an understanding of human behavior. At least part of the lesson to be learned from these events, it would seem, is that social forces and environmental contexts may in certain instances be prepotent over individual traits or may in any event have such immense implications and impact that they cannot be ignored without penalty in any analysis of individual behavior.

We are not without warning in this respect. Sullivan (1953, 1956), Lewin (1935, 1936), Fromm (1941, 1955), Murray (1938), Murray and Kluckholm (1953), and Getzels and Thelen (1960) have all presented theoretical points of view that emphasized the importance of environmental and social determinants and their interactions with individual needs and characteristics. Gage (1963) reviewed paradigms for research on teaching, several of which incorporate concepts related to this kind of interaction. Would that research methodology could only implement all of them properly! Inkeles (1963) proposed that social action responses can be

conceived fruitfully as the resultant of social conditions and personality needs and their interactions; he also provided some useful suggestions for the articulation of methods and findings from sociology and psychology. The assessment procedures of Stern, Stein, and Bloom (1956) emphasized the importance of analyzing role requirements in terms of the requisite needs or personality characteristics of the individual most likely to perform in those roles. Pervin (1968) reviewed research relating to "Performance and Satisfaction as a Function of Individual-Environment Fit." But in spite of such evidence, the current situation is much the same as it was in 1955 when Julian Rotter (1955, p. 247) made the following statement:

In the half century or more that psychologists have been interested in predicting the behavior of human beings in complex social situations they have persistently avoided the incontrovertible importance of the specific situation on behavior. They have assumed that if they could only produce a somewhat better schema for attempting to describe an individual's personality from a purely internal point of view they could somehow or other overcome this failure to predict. So they have gone from faculties and instincts and sentiments to traits, drives, needs, and the interaction of these within the individual, producing schema for personality organization and classification of internal states, but ignoring an analysis of the psychological situations in which human beings behave.

Rotter reported in the same article (p. 246) that after he had conducted an exhaustive search of the literature for systematic schema to be used in predicting behavior from specific situations, he became convinced that "... psychological climate is like weather; everyone talks about it but no one does anything about it."

It may be high time to do something about it. If psychology can not be more effective in contributing to the understanding and solution of the major educational and social problems of the day, it may be so subject to certain distortions and perversions of its fundamental intent that it loses any claim to objectivity as a science. Maslow's (1969) description of a particular brand of "third-force" psychology as "introspective biology" may have appeal for some but may constitute something less than a "peak experience" for others. As it becomes less certain that psychological schema interpret or predict anything, there is an increasing tendency to develop assumptive structures that create their own realities rather than attempt to define, explain, and predict the seemingly intractable reality that confronts us in everyday life. Some become increasingly philosophical and singleminded in their orientation and tend to stray from the monumental but essential empirical tasks that lie before them. *A la*



Kelly (1955), their psychological processes become channelized by the ways in which they anticipate events, and the channels grow deeper and deeper.

If the person-environment interaction is critical for understanding and predicting human behavior, it is equally apparent that this interaction can only be defined effectively in multivariate terms. We are multi-trait individuals responding to multi-characteristic environments, and the total pattern of these interactions determines the direction of our behavior. If this is accepted as fundamental, the research methodology appropriate to such a conception becomes a major issue. The work of Campbell and Stanley (1963) on "Experimental and Quasi-Experimental Designs in Research on Teaching" has become a classic, and most deservedly so, but a strong insistence on variable-manipulating studies may not always be compatible with either a person-environment model or the multivariate approach it seems to necessitate. Cronbach (1957) pointed out some of the issues here in his discussion of differing approaches of the "experimental" and "correlational" psychologist and saw hope for a productive rapprochement in the future. Whether psychology has come any closer to that rapprochement in the twelve years since the publication of Cronbach's manuscript is an issue worthy of conjecture.

It would seem that a pertinent set of criteria for a research methodology in psychology that would have greater potential for advancing the field and contributing more efficiently in the long run to the solution of our critical educational and social problems would include the following:

1. Conceptualization of research problems, as often as feasible, within the framework of person-environment interactional systems;
2. Conceptualization of research problems in multivariate terms that accurately reflect the complexity of both personological and environmental domains;
3. Provision of appropriate techniques for measuring critical environmental and situational variables that are at least as reliable, valid, and precise as those techniques we now employ for measuring personological variables (Cronbach, 1957, p. 677);
4. Provision of methods for defining and assessing person-environment relationships that are theoretically, logically, and operationally defensible;
5. Application of multivariate statistical methods appropriate for analyzing the data derived from such a research setting.

The remainder of this paper is devoted to the review of a few selected research efforts that involve person-environment interaction of some kind and which may be analyzed and evaluated in terms of the criteria above. The careful consideration of this research in relation to these criteria can

further sharpen our perception of needed directions for future research and the benefits and problems that will probably be encountered if psychologists move in the directions prescribed. Since psychology is such a broad field, even in its application to education, this review requires extreme selectivity and focus. A unifying theme is required. The unifying theme of this review is that if educational psychology hopes to improve its predictive efficacy and utility, it must now make a strong effort to develop scientifically sound methodological approaches adequate to the task of making accurate predictions of individual behavior from the context of empirically assessed individual-environment interactions. The research areas reviewed are those that, in the author's judgment, are interesting and illustrative examples of systematic methodological approaches employed in selected current research efforts involving some kind of person-environment interactional framework. Because of the very nature of this assignment these examples are far from exhaustive and do not include many investigations involving naturalistic and relatively subjective observational settings and many others that are of an *ad hoc* nature and less classifiable in terms of methodological approach. The assignment demands selectivity and precludes representativeness, and it is important that the reader keep this in mind throughout the review.

It is also true that substantively the examples provided may not always be directly related to the contemporary social issues referred to earlier, since changes or improvements in research technique are not always made in the areas of most crucial social need. But methodologically they constitute some of the best available examples of a generalizable research orientation, always involving person-environment interactions but with many methodological variants, that should be fully explored to determine its potential for improving the predictive efficacy of our behavioral science and contributing ultimately and efficiently, it is hoped, to the understanding of the human being *in situ* and his social ills. If this orientation is critical to the progress of psychology as a science, particularly in its application to education and society, the solution of any methodological problems that may be associated with it is a necessary condition of that progress. In the reviews that follow various research efforts representing that general research orientation are considered along with the methodological issues associated with each. Selective reviews are provided of three current but distinctively different areas of research involving some kind of person-environment interactional framework: aptitude-method interactions, verbal interaction analysis, and person-environment interaction on the college campus.

### *Aptitude-Method Interactions*

Any consideration of research related to person-environment interactions must acknowledge that what the researcher has chosen to isolate

as "environment" is essentially arbitrary and is dictated by his personal interests, the nature and the hypotheses of his study, the necessity for organization and focus, and his awareness that presently available research techniques place some practical quantitative restrictions on his domain of operation. For several years Cronbach has been interested in the possible interaction of individual aptitudes with that particular class of environmental variables defined as instructional methods or treatments. In an early statement Cronbach (1957, p. 680) placed himself firmly in the interactionist camp by asserting that:

Applied psychologists should deal with treatments and persons simultaneously. Treatments are characterized by many dimensions; so are persons. The two sets of dimensions together determine a payoff surface. For any practical problem, there is some best group of treatments to use and some best allocation of persons to treatments. We can expect some attributes of persons to have strong interactions with treatment variables. These attributes have far greater practical importance than the attributes which have little or no interaction.

In a later paper Cronbach (1967) called for research which would help establish which kinds of instructional methods have most achievement payoff for which kinds of pupils. His call was for a uniquely different kind of research on instruction, an approach which would forgo the assumedly fruitless search for the one best instructional method and would seek instead the most effective methods for different kinds and levels of aptitudes. This research would furnish the theoretical and empirical foundation for the systematic adaptation of methods to pupil differences, and Cronbach suggested the direction that this effort should take:

To systematize the process of adaptation, and hence reduce error, calls for a theory whose propositions would state the conditions of instruction best for pupils of certain types, both conditions and types being described in terms of fairly broad dimensions (cf. Jensen, 1962). I doubt that we can develop a separate rationale for each area and level of curriculum, but I have hopes for the more general theory applicable at many levels and in many areas. Such a theory deals with aptitude-treatment interactions (Cronbach, 1967, p. 30).

Cronbach cited little research evidence relevant to this point of view simply because there is little useful evidence of this type available. His plea was for more researchers to acknowledge the complexity and accept the challenge of research in this area. He also cautioned about the possibility that teachers might *over-differentiate* on the basis of differential information (i.e., place greater weight on pupil differences and conse-

quently provide a greater variety of instructional methods than would be justified to achieve maximum achievement within the classroom group). Although at first glance this seems unlikely, Cronbach (1955, pp. 182-183) cited some related evidence that counselors were guilty of a similar kind of overdifferentiation when asked to predict student grade average from data provided them. In this case the counselors placed more weight on the predictors than was justified and made some extreme predictions. But in spite of such pitfalls it is clear that Cronbach sees the advantages of research on aptitude-treatment interactions far outweighing the risks involved.

Research is always a gamble, and it is difficult to predict the payoff from such an approach. In commenting on Cronbach's paper, Carroll (1967) agreed that the study of instructional methods and individual differences would be psychologically interesting but might be extremely difficult and frustrating as well; he suggested that "reality testing" in this area might be painful. He pointed out that research might not be able to produce a sufficiently solid set of conclusions to justify being adopted in educational practice, or that the cost of differentiating instruction in accordance with research findings might be too high in relation to the increase in efficiency obtained.

Bracht and Glass (1968) issued some additional caveats. In a paper focusing on problems of external validity in experimentation, Bracht and Glass listed the interaction of personological and treatment effects as one of the "threats" to the external validity of experiments. Their orientation here serves to recall Cronbach's earlier assertion that the experimentally oriented psychologist tends to see individual differences as an annoyance rather than a challenge; since his goal is to control behavior, such variations within treatments are proof that he has not succeeded (Cronbach, 1957, p. 674). One can also see a parallel with Spearman's (1927) disposition to regard the residuals from the tetrad difference equation as "disturbers" in the context of research on the general intellective factor. Cronbach was right; what is one man's "disturber" is another man's "peak experience." In any event, Bracht and Glass did make some very useful points about aptitude-method interactions that are relevant to the present discussion. They asserted that what little evidence there is for personological variable by treatment interactions is, in their judgment, not very convincing. They also presented some important considerations related to whether evidence of significant interactions justifies differential treatment of persons. They pointed out that a significant interaction does not necessarily imply the crossing of regression lines, and that a significant interaction *and* the crossing of regression lines do not necessarily justify differential treatment of subjects differing on the personological variable. They discussed the distinction between ordinal and disordinal interactions and indicated the need for inferential statistical methods to determine



whether the crossing of sample regression lines in the latter case is sufficiently great to suggest disordinal interactions in the population parameters that would justify differential treatments.

Clearly, there is a challenging and rocky road ahead and no assurance that the destination will be reached or that the yield will justify the effort. But at this stage in the development of educational psychology it is important that the effort be made, and the likelihood of payoff would probably be greater if it were made within the context of a large-scale research program of the type that could only be promoted by a research and development center or similar organization. If legitimate and psychologically significant aptitude-method interactions were discovered, they could have considerable influence on our educational planning, especially for the disadvantaged. The possible social significance of such findings has been underscored by Jensen (1968, p. 23):

Perhaps our greatest hope of achieving equality of educational opportunity lies in the possibility of finding significant *patterns* of individual differences in the development of abilities and in taking advantage of these differences to create the optimal Instruction pupil interaction. We have seen evidence that this can happen in the learning laboratory (Jensen, 1967). *If it is a false hope for school learning, we can find this out by making a serious attempt.* (Italics mine)

One useful distinction that could be made about ability patterns, according to Jensen (1969, p. 110), is that contrasting "associative" and "conceptual" abilities. He cited evidence of their disproportionate representation in different socioeconomic groups and their different developmental characteristics within such groups. The implications he drew for instructional practices are clear—as is the potential of such implications for engendering political irascibility.

In the face of the possible political overtones that may accompany findings on aptitude-method interactions, the challenge to research methodology assumes an added dimension. There must be a very high degree of confirmation for such interactions before there can be any thought of their implementation in instructional practices. Bracht and Glass (1968) contend that the likelihood of such confirmation for the more molar variables (like "associative" and "conceptual" abilities) is much less than when both variables and treatments are more narrowly defined. Herein lies another possible dilemma for interaction research. In the statement by Cronbach quoted earlier in this chapter, Cronbach stressed the probable difficulty of developing a separate theoretical rationale for each area and level of curriculum and suggested instead the specification of a theoretical rationale involving broadly defined instructional treatments and pupil types as the best approach to studying aptitude-treatment inter-

actions. Yet these are the very kinds of variables least likely to reveal the interactions sought, according to Bracht and Glass. According to these authors (Bracht and Glass, 1968, p. 452), any search for such interactions "with treatments as necessarily complex as instructional curricula," for example, "may be fruitless." Evidence for this assertion is presented in a recent dissertation by Bracht (1969).

It seems also that the interactions sought are sometimes accorded a kind of timelessness they do not deserve. A given interaction may be more time-bound to a particular level of development than is commonly recognized. Many interactions may be in fact three-way interactions, with the temporal dimension as the third factor. Upon reaching a certain level of development, certain interactions may emerge and others disappear. Piaget's concepts relative to *adaptation* would seem to be relevant here. According to Piaget (1952), adaptation to any cognitive task involves two processes: *assimilation* and *accommodation*. *Assimilation* involves the structuring or restructuring of an object to make it congruent with the cognitive structure of the organism. *Accommodation* is the reciprocal process involving changes in the cognitive structure to facilitate the assimilation of the object (learning, cognitive task) in question. Thus, every cognitive adaptation involves some change in the cognitive structure of the organism. It is conceivable that a particular history of accommodative acts and the presence or absence of particular aptitude-method interactions are closely interrelated, and that the emergence and disappearance of these interactions in relation to these accommodative processes is more rapid than might initially be thought.

This mixing of qualitatively different streams of theoretical thought may seem as bad in its effects as the mixed metaphor, but it does serve to alert researchers to the possible importance of considering the developmental dimension in studies of aptitude-method interactions. Indeed, some interactions may prove to be so ephemeral in a developmental sense that they might not be worth studying. And these may also involve the very same "narrowly defined" variables that Bracht and Glass felt were most likely to reveal the interactions that are sought. In any event, it would seem that the developmental factor should be very carefully considered in any research methodology adopted for the study of aptitude-method interactions.

Finally, research on aptitude-method interactions can not be conducted as if the behavioral payoff were only singular. There are many different kinds of learning, all with their own performance criteria, and there is no assurance that aptitude-method interactions appearing with the application of one criterion or set of criteria would appear with another, or that interactions would be characteristically the same if they did appear. Furthermore, the likelihood of dissimilar findings with respect to interactions probably increases with the extension of performance or

behavioral criteria beyond those encompassed by the behaviors usually labeled as learning. Interactions that might be used to maximize learning payoffs, for example, may be quite different from those that might be used to maximize student satisfaction, or motivation for additional study, or the drawing of implications for one's personal behavior. In effect such studies involve not two-way interactions but at least three-way interactions, viz., aptitude-method-performance criteria interactions, and this makes the task enormously more complicated.

As an example of a research endeavor that focuses on one type of person-environment interaction, research on aptitude-method interactions is obviously not without its problems. In terms of the five criteria that were earlier specified it seems to meet only the first and fourth criteria with any degree of certainty: it provides a most seminal conceptualization of the possible importance and effects of these kinds of interactions, and it usually casts its research problems in terms of person-environment relationships that are logically, theoretically, and empirically sound. The other three criteria are not met nearly as well, probably because of the relative infancy of research effort in this area. Conceptualization in multivariate terms is not as likely until the results of simpler investigations are in and evidence begins to accumulate that the approach is fruitful. Techniques for measuring the critical environmental (method or treatment) variables are at present often global and imprecise, especially when dealing with the broad and complex categories of instructional method. Until further advances in these areas are made, the fifth criterion (application of appropriate multivariate statistical methods) is not entirely relevant to the present state of the art. Much remains to be done. Whether it proves to be worth the effort can only be ascertained after the effort has been made and the data are in, but the effort must certainly be made. Pioneering efforts in this direction have already been made by Bracht (1969), Di Vesta (1969), and Cronbach and Snow (1969). Results of the Di Vesta project have not yet been reported. Bracht's research yielded a second-order ordinal interaction suggesting differential effects of "verbal" and "spatial" methods of teaching the addition of positive and negative numbers to pupils of different performance levels on verbal and spatial ability tests. The Cronbach and Snow report (1969) indicated that few or no aptitude-treatment interaction effects have been solidly demonstrated, either in their research or the research of others, but advocated further commitment to what is admittedly high-risk research and offered helpful counsel about the direction that this research should take. These are all useful resources to anyone planning a research study in this area.

### *Verbal Interaction Analysis*

Current examples of research on person-environment interactions come from research settings that are quite different from one another in

terms of conception, development, and results. Another class of environmental variables that has excited some interest is that broadly referred to as the social climate of the classroom. Withall and Lewis (1963) and Medley and Mitzel (1963, p. 263) provided comprehensive reviews of research on social interaction in the classroom and classroom climate. More recently Meux (1967) discussed studies of learning in the school setting and recent developments in classroom observation systems. It is not appropriate, either in terms of space limitations or present objectives, to present such a review here. There is, however, one current example of research in this area that has stimulated much interest and attention not only because of its implications for research but also because of its widespread and often indiscriminate extension to teacher education programs. This is the system of verbal "Interaction Analysis" espoused by Flanders et al., and it is to this kind of analysis that we direct our attention in the present section (see Amidon and Hough, 1967).

The type of interaction analysis proposed by Flanders can be characterized as content-free and also method-free, depending on the inclusiveness of one's concept of method. In any event, it emphasizes quite a different set of variables than Cronbach apparently had in mind when he referred to instructional method. Flanders developed a system of observing and recording classroom behavior in terms of the verbal interactions of the teacher with students. The observer makes judgments about communication behavior every three seconds and classifies that behavior into one of ten categories: (1) accepts feelings, (2) praises or encourages, (3) accepts or uses ideas of student, (4) asks questions, (5) lectures, (6) gives directions, (7) criticizes or justifies authority—all of which are classified as "teacher talk," and (8) student talk-response, (9) student talk-initiation, and (10) silence or confusion (Amidon and Flanders, 1963). The numbers corresponding to each three-second segment are recorded in sequence and subsequently entered into a data matrix. The data matrix is a 10 by 10 table corresponding to the observational categories; each observational judgment is recorded by number in the appropriate column and in the row corresponding to the category number of the observational judgment immediately preceding it in time. This ingenious recording system permits the analysis of the antecedents of verbal acts and possible cause-effect relationships. Because of this feature and the division of categories into "teacher-talk" and "student-talk," it is possible from matrix data to analyze segments of behavior that are truly interactional in nature. The interaction is predominantly of a teacher-group nature; the potential for studying teacher-individual student interactions is limited without further elaboration and revision of the technique.

Several methods for analyzing the data matrix are possible. An initial step is to compute the total percentage of teacher talk, the total percentage of student talk, and the percentage of each of these that is devoted



to each of the singular categories of behavior. Two important scores are the ID and revised ID ratios. The ID ratio is the ratio of indirect to direct teacher statements and is defined by the total number of tallies for categories 1, 2, 3, and 4 divided by the total number of tallies for categories 5, 6, and 7. The revised ID ratio (*id*) was devised to represent an indirect-direct ratio that omitted considerations of subject matter presentation and reflected the kind of emphasis given to motivation and control in the classroom; it is defined as the total number of tallies for categories 1, 2, and 3 divided by the total number of tallies for categories 6 and 7. The emphasis given to these ID ratios reflects Flanders's judgment, with considerable empirical support, that the teacher's verbal behavior is the principal determinant of the classroom environment. In addition to these ID ratios, certain sections of the data matrix are also singled out for specialized kinds of analysis.

Interaction analysis is a striking example of a technique that may be ruined by its own success. It is a most ingenious technique with some very commendable features, and its more enthusiastic devotees do not hesitate to describe it as a "breakthrough" that will be a bonanza for teacher education (Campbell and Baines, 1969). The present author does not want to diminish in any way the contribution that interaction analysis has made and will continue to make; he is well aware that it is much easier to criticize than to create. The technique is a most fruitful step in the right direction, and whatever is said in the following paragraphs should be interpreted in the light of this recognition of its over-all contribution.

As a research methodology, however, there are some aspects of interaction analysis that suggest the need for further reflective thinking and additional research instead of a head-on attempt to convert the heathen. Concerns develop initially when one examines the "ground rules" that are given to those who intend to use interaction analysis for observing and recording classroom behavior. Rule 2, for example, states that:

*If the primary tone of the teacher's behavior has been consistently direct or consistently indirect, do not shift into the opposite classification unless a clear indication of shift is given by the teacher* (Amidon and Flanders, 1963, p. 21).

In expanding upon this point it is indicated that the observer is reacting to the *general tone* of the teacher's influence, either direct or indirect, and does not use the opposing categories unless it is clear that the teacher has shifted from his more general pattern. The rule is also called the "rule of the biased unbiased observer," suggesting again that the initial bias of the observer, based upon the aforementioned general tone of the teacher's influence, should not be altered unless there is good reason to do so. Confronted with such a rule, one cannot help wondering what this might mean for independence of observations and the possible develop-

ment of halo effect—either positive or negative. Are observers making independent observations or developing biases about the directness or indirectness of a given teacher's behavior? And if such biases are developed, what is to prevent them from emerging as a result of the interplay of certain unconscious criteria in the observer's mind with the supposedly focal "direct" and "indirect" criteria? This ground rule seems to call into question the objectivity of the entire observational process and to suggest that the ten variables being rated may sometimes be *post hoc* derivations from generalized biases which may or may not stem from the constructs of "direct" or "indirect" teaching.

Rule 3 of the series of ground rules also has some interesting implications:

*The observer must not be overly concerned with his own biases or with the teacher's intent. Rather he must ask himself the question, "What does this behavior mean to the pupils as far as restriction or expansion of their freedom is concerned?"* (Amidon and Flanders, 1963, p. 22.)

The last sentence prescribes a focus for the observer, but it might well be asked how fruitful such a focus is or whether it captures even a small part of the complexity of classroom interaction. Are we not now past the stage when we can be content with simplistic molar variables of the democratic-undemocratic and good guy-bad guy variety? The direct-indirect and expanding-restrictive constructs are at about the same level of conceptualization as dominative-integrative (Anderson and Brewer, 1945; Anderson, Brewer and Reed, 1946), teacher-centered vs. learner-centered (Withall, 1949), and hostile-supportive (Medley and Mitzel, 1958) constructs that have been used in previous studies. Although reeking with value judgments, these constructs have been useful in the past to stimulate thinking about classroom interaction and assessment and to get started on a difficult research endeavor, but continued pursuit of research at this level of conceptualization is not likely to result in much progress in understanding and predicting classroom interaction patterns and their correlates. Both observation and the complexity and confusion of research results on teacher effectiveness seem to suggest that there are many different *patterns* of variables that can be associated with effective teaching and with ineffective teaching; these are not likely to be easily subsumed under categories like "direct" and "indirect" teaching without some damage to the data and the reality that the data reflect. Research methods must be conceptually and statistically multivariate to a far greater degree than they are now, and with far greater sensitivity to the nuances of teaching style and pattern than they are now, before much further progress can be expected.

The foregoing comments also lead to other considerations. With molar

variables and research methods of the type described, there is more likelihood of interpretive error and the neglect of alternative cause-effect explanations. It is contended by Flanders (1964), for example, that indirect teaching is responsible for more positive pupil attitudes, with very clear and specific indications that the former is seen as cause and the latter as effect. Might not the cause-effect relationship be quite different from this? Might it not be that poor teaching, of whatever variety or cause, creates student dissatisfaction and misbehavior, and that a teacher perceiving this will become defensive and resort more frequently to lecture, giving directions, criticizing, or justifying authority—all of which are “direct” variables in interaction nosology? Rather than direct teacher influence causing negative pupil attitudes and behavior, negative pupil attitudes from poor teaching of any kind would be seen as causing direct (defensive) teacher behavior. There is some evidence suggestive of this alternative explanation. In one study, for example, it was found that direct teachers “. . . had more discipline problems and found it necessary to interrupt giving directions in order to criticize students three times more often than did the most indirect.” (Flanders, 1964, p. 210). Is such a strong effect more likely as a product of merely “direct” teacher behavior or as a concomitant of markedly deteriorated pupil behavior attributable to many kinds of grossly ineffective teaching practices? It is also interesting that in another study dependent-prone students learned more geometry from indirect than direct teachers, although theory might have predicted otherwise (Amidon and Flanders, 1961). This finding becomes more explicable in terms of the alternative rationale suggested above. If the direct teacher category merely includes many varieties of ineffective teachers teaching defensively (i.e., directly), it can be concluded that no student, of whatever personality characteristics, could be expected to learn much from these ineffective teachers. The issue of alternative cause-effect explanations can not be decided on the basis of presently available evidence, but there are enough questions about causation here to justify research methods more sensitive to possible alternative explanations.

These issues are not theoretical issues alone, since teacher education programs are now being built upon the findings and results from interaction analysis. In some reports statements like the following are made:

The scientific study of teaching is so immature that, at this time, a particular pattern of teaching cannot be advocated as the most successful (Flanders, 1964, p. 229).

. . . each teacher needs the freedom to develop his own unique style of teaching (Flanders, 1964, p. 230).

while in other instances statements seemingly inconsistent with these are also being offered:

The purpose of the in-service training was to increase the flexibility of teacher influence and to increase the use of those teacher behaviors which support pupil participation in the classroom learning activities (Flanders, 1963, pp. 27, 28).

It is not to be denied that the classroom performance of many teachers would be immensely improved if they shut up on occasion, but the creation of teacher education and in-service programs focused relentlessly on the development of one kind of teaching pattern seems little justified by present research findings, poorly advised, and inconsistent with previous utterances about the state of the art.

Some of the hypotheses that are developed within the framework of interaction analysis seem to lack a convincing theoretical rationale and often take on the appearance of an *ex post facto* structuring of a *potpourri* of confusing results. Three hypotheses that have been proposed, for example, are that (1) restricting student freedom of participation early in the cycle of classroom learning activities increases dependence and decreases achievement; (2) restricting student freedom of participation later in the cycle of classroom learning activities does not increase dependence but does increase achievement; and (3) expanding student freedom of participation early in the cycle of classroom learning activities decreases dependence and increases achievement (Flanders, 1964, p. 202). The indirect teachers are supposed to be more flexible than the direct in their ability to shift from expanding to restricting activities, or vice versa, at the appropriate times (Flanders, 1964, pp. 215-219). Although these hypotheses are not altogether congruent with common sense, they could be accepted on the basis of strong evidence for their support. However, the research evidence is often mixed, only weakly supportive, or uncertain, and often based on a very limited number of teachers. Histograms are often offered as evidence in the place of conventional and more convincing data presentation. But despite the absence of a convincing theoretical rationale and conclusive research evidence, the three hypotheses stated above are still embedded in ongoing programs of teacher education. In a heuristic sense this may not be regrettable; any program that stimulates teachers to be self-analytical may do more good than harm. But from the standpoint of research methodology it suggests the need for more exacting criteria and an extended program of research.

Any such program of research should include some additional attention to the reliability of rating categories. In a report intended for researchers, Flanders (1965, p. 30) made the following comment:

Even a casual inspection . . . gives ample evidence that error in a single category, on a particular observation, can be very high when the frequency of events in that category is low. Interaction



analysis is far from a precision instrument. As a research tool, it should be used cautiously.

Such a comment is most laudable, but it is also puzzling in relation to past and current attempts at wholesale application of the technique to teacher education programs. Many reliability problems remain to be worked out. The reliability figures reported are based on total data matrices and do not represent interrater reliability for subscores and subsections of the matrix that are employed in recommended data analyses. They are also based on the incorrect application of Scott's (1955) coefficient. There is evidence that the error rate is at its greatest for those categories that contribute most to the differentiation of "direct" and "indirect" teachers (see Amidon and Hough, 1967, p. 138, 164), and the difference between these two types of teachers in percentage terms is quite small (see Flanders, 1964, p. 215). Furthermore, the reliability coefficients that are provided are based on cumulative errors and not absolute errors, i.e., they represent percentage comparisons of the same categories for different raters, but not the degree of agreement between two raters on exactly the same observed events (Flanders, 1965).

In terms of the five criteria that were earlier suggested for measuring progress in developing research methods that would be valid for studying person-environment interactions, it appears that interaction analysis meets the first criterion and part of the third criterion reasonably well. Interaction analysis provides a conceptual structure for analyzing certain kinds of interactions, and it provides an ingenious set of observational and analytical tools for interpreting teacher-determined environment. Problems remain, however, with the objectivity, validity, and reliability of the ratings and the various analytic operations. The fourth criterion is also met reasonably well, although reservations about reliability and objectivity apply with equal force to questions of person-environment relationships. Alternative cause-effect explanations of person-environment relationships should also be considered more often than they are. Interaction analysis is found wanting with respect to the second criterion, for its major constructs appear too limited in scope to represent the complexity of the classroom situation. And in its application of multivariate statistical methods (criterion #5), it sometimes appears to be too fixated at the simplistic "indirect-direct" conceptual level to take adequate statistical advantage of the data that are available or could be made available. But if problems have been encountered, it is because the task is so difficult, and encouragement should be shouted to those who are willing to undertake the challenges involved. As methodological problems are solved, and as additional evidence is amassed, the contributions to the understanding of individual-environment interactions in the classroom should be significant indeed.

## *Person-Environment Interaction on the College Campus*

The environmental aspect of what has been referred to as person-environment interactions has been abstracted and conceptualized in quite different ways in the two preceding sections. For the purposes of the first discussion the environment was defined as the instructional treatment or method presented to the student. For the immediately preceding discussion it was defined as the verbal behavior of the teacher in terms of its influence on and interactions with pupil behavior. The teacher's verbal behavior, according to Flanders, is the principal determinant of the kind of environment that exists in the classroom. The research area now to be considered defines environment in terms of a much more comprehensive setting, that of the total school environment. Thus defined, "environment" comes closest to what Sells (1963, p. 700) had in mind when he pleaded for an interactionist point of view in his address "An Interactionist Looks at the Environment":

The most obvious need in evaluating the manifold encounter of organism and environment is a more satisfactory and systematic conceptualization of the environment. This implies a taxonomic, dimensional analysis of stimulus variables comparable to the trait systems that have been developed for individual difference variables.

In the same address Sells lamented the fact that:

While work proceeds actively to extend the exploration of individual differences, however, the equally important frontier of situational dimensions is virtually ignored.

To say that such situational analyses have been virtually ignored is quite true, and yet some progress has been made in this respect, particularly since the late 1950's, in the development of conceptual schemes for characterizing the environment of the college campus. In this section the available evidence on person-environment interactions on the college campus will be considered. Because of the necessarily selective nature of this review, only that work concerned with the college environment will be considered and not research on organizational climate, the high school environment, or other related research.

Of the different techniques available for the assessment of the college environment (Astin and Holland, 1961; Astin, 1963b; Creager and Astin, 1968; Nunally, Thistlethwaite, and Wolfe, 1963; Pace, 1963; Pace and Stern, 1958; Werts, 1967), the most widely used are classifiable into two general types. The first type, represented by the College Characteristics Index (CCI) (Pace and Stern, 1958) and the College and University Environment Scales (CUES) (Pace, 1963), requires students to indicate

whether certain statements are characteristic or not characteristic of their college environment as they perceive it. The second type was developed partly as an effort to overcome the felt shortcomings of the first and avoids the student perception or "image" approach to environmental assessment; in the case of the Environmental Assessment Technique (Astin and Holland, 1961), for example, assessment is made of eight objectively determined indices: size and intellectual level of the student body, and the percentage of the total student body awarded degrees in programs corresponding to each of six major areas (realistic, intellectual, social, conventional, enterprising, and artistic). Both types of instrument have been used widely in studies comparing college environments, but these studies will not be reviewed here because they are peripheral to the present focus on person-environment interactions (Astin, 1962a, 1965a; Pace, 1960; Stern, 1962a, b; 1963a). Readers are referred to the review of "Campus Environment" in the October 1965 issue of the *Review of Educational Research* (Michael and Boyer, 1965).

One troublesome and confounding kind of person-environment interaction which constitutes a serious methodological hurdle for instruments of the CCI and CUES type is that student perceptions of their college environment may be influenced and distorted by their own personality characteristics. In an early study McFee (1961) concluded that there was no relationship between student needs and their perceptions of the college environment, but some later studies revealed such relationships (Baker, 1966; Herr, 1965; Marks, 1968; Mitchell, 1968; Stricker, 1967). For studies employing a phenomenological frame of reference this might be considered useful or advantageous in some respects, but for studies of person-environment interactions it serves to confuse person variance with environment variance and is likely to lead to uncertain conclusions. This is particularly true if the personal trait of nonconformity tends to be related to environmental perceptions, as Mitchell (1968) found in a recent study.

Another kind of confounding interaction likely to lead to interpretive confusion is revealed by Stern's finding that colleges with certain environmental press characteristics tend to attract students whose need patterns are personalized versions of the prevailing press (Stern, 1960a, b, 1962a, b). (The evidence would seem to indicate that this effect exists separately and in addition to the effect noted in the paragraph above.) Stern's finding achieves even greater significance in relation to Astin and Holland's (1961, p. 315) finding, in validating the Environmental Assessment Technique, that EAT-CCI correlations were sufficiently high to suggest that "... the attributes of the student body reflect a major portion of what has been called the college press or environment." It thus appears that in studies involving the analysis of environmental effects on students, it is imperative that the variance attributable to student input character-



istics be taken into account before certain student behaviors are attributed to the effects of the college environment.

The importance of the latter consideration was revealed in a series of early research studies dealing with the possible influence of the college environment in stimulating undergraduates to seek the Ph.D. degree. In two early studies Knapp and Goodrich (1952) and Knapp and Greenbaum (1953) presented evidence that the "productivity" of undergraduate institutions, defined as the percentage of undergraduates who eventually obtained the Ph.D., differed widely from institution to institution, and attributed these differences to differential environmental influences. Holland (1957) then pointed out an interpretive problem: since these colleges also differed widely in the intellectual quality of their entering students, one could not be certain whether these differences in productivity arose from differences in environmental influences or differences in initial student ability. Thistlethwaite (1959), using the same criterion, avoided the earlier error and partialled out the effects of student quality. He reported significant and relatively high correlations of productivity with CCI press scales. Sophistication in research methodology was increasing. It increased still more when Astin (1961) showed that student ability was not the only input that had to be partialled out. He correlated the productivity criterion (with ability partialled out) with the percentage of students in each college who aspired to the Ph.D. upon entering and the percentage who planned to major in a natural science field at the time of entrance; the correlations were .68 and .60 respectively. After partialing out these additional two variables, the correlations between CCI press and productivity were reduced considerably. Press variables like "Reflectiveness" and "Energy and Controversiality of Instruction," however, still correlated significantly at the .05 level. More recently some new methodological issues were raised with the contention that part correlation of the type commonly used in school environment studies to control the effects of student input variables may lead to spurious estimates of environmental effects (Werts, 1968), depending on the model used (Werts and Watley, 1968). Alternative procedures were suggested by the authors.

For those interested in the challenge of thorny methodological problems, this area is a real bonanza. A spirited exchange of methodological viewpoints resulted from one study (Thistlethwaite, 1962a; Astin, 1962b; Thistlethwaite, 1962b). Problems of control, especially for student input variables, loom large, and the investigator is seldom certain that his methods have assured statistical control of the critical input variables so as to allow relatively secure inference about direction of causation from correlational data. But methodological sophistication has increased with exposure to the pitfalls, and this progress has been made over a relatively short period of time. There is now some evidence of varying strength that Ph.D. productivity is related to certain environmental press variables



(e.g., enthusiasm, humanism, affiliation, independence, achievement, and supportiveness) (Thistlethwaite, 1962a, b); that *changes* in plans to seek advanced training are positively correlated with excellence of faculty in major field and negatively with faculty press for compliance (Thistlethwaite and Wheeler, 1966); that career choice at graduation is affected by the proportion of other students in the student body planning careers of a particular type (Astin, 1965b); and that withdrawal from college is associated with certain environmental press, e.g., lack of concern for the individual student (Panos and Astin, 1968). But after control of relevant student input variables, the relationships revealed are of a rather low order of magnitude, and the state of affairs has prompted one leading researcher in the field (Astin, 1963a, pp. 70, 71) to assert that:

The difficulty encountered in interpreting these significant college effects indicates that a more intensive study of student-college environment interactions is necessary. The gross college variables used in this study (i.e., the EAT and the factored dimensions) should continue to be useful in identifying differential college effects, but a more meaningful interpretation of such effects will depend ultimately on our ability to describe the college environment in terms of the experiences of the individual student.

In reality, however, research studies of the type reviewed in the last paragraph are not truly interactional in nature, since they are *not* concerned with differential environmental effects on students of different personality characteristics. That there are very few studies of this type is not at all surprising in view of the many methodological problems involved. There are serious problems, for example, in identifying, conceptualizing, and measuring the environmental variables that have most influence on and interactions with individual behavior. Murray (1938) distinguished *alpha* environmental press from *beta* press, the former being environmental objects as they exist in reality or as objective inquiry discloses them, and the latter as they are perceived and interpreted by the individual. The CCI and CUES are based on the latter rationale, and the EAT attempts to define empirical indices of the former. Even though based on different rationales, these two different kinds of instrument tend to correlate with one another (Astin and Holland, 1961), but neither kind is very predictive of individual behavior. Perhaps Astin is correct in the quotation above; these global variables may be too far removed from the experiences of the individual student. The environmental variables that may count the most may be those that impinge directly on the individual through his dyadic or small group interactions. These may be the first-order effects, with the second-order effects of the larger environment being influential only insofar as they broadly define the terms of the milieu and the kinds of first-order interaction that are generally con-

gruent with it. The phenomenological frame of reference is inescapable, and the individual's interaction with the environment in terms of his personalized interpretations of environmental events may be the critical methodological challenge. And the largely adventitious nature of many of our individual experiences will not make the problem of prediction any easier.

Another difficult methodological problem in this area is that of developing a reasonable theoretical rationale and methodological implementation for defining and measuring person-environment interactions and predicting their effects. Both the CCI and the EAT are based on interaction rationales, but they have led to few studies in which the empirical implementation of that rationale met the promise of the original conceptualization. The CCI is based upon Murray's conception of need-press interactions or *themas*; the individual's behavior was thought to be best predicted from a knowledge of how the relevant environmental press variables facilitated or impeded the satisfaction of an associated need (Murray, 1938). The CCI "press" scores have their counterpart "need" scores in the Stern Activities Index (Stern, 1963b), which permits an analysis of institutional press in relation to the need pattern the student body brings to the school. Stern (1960a, 1962a) suggested this as the sociopsychological equivalent of an ecological analysis, but it tends to fall short of this in actuality. Most of the evidence of this nature provided so far has been of an intuitive-descriptive variety or has involved the comparison of standardized press scores with their counterpart student body need scores (Stern, 1960a). The need-press comparisons often pose tricky theoretical questions. What, developmentally and psychologically, is the optimal press counterpart to a need for aggression or abasement or deference or fantasied achievement? Neither the psychological rationale nor the normative standards exist for determining need-press "goodness of fit" for individuals or groups.

The Environmental Assessment Technique (Astin and Holland, 1961) is similarly based on an interaction rationale, in this case a prediction model for vocational choice (Holland, 1966). Both people and environments are categorized as one of six types: realistic, intellectual, social, conventional, enterprising, and artistic. Regarded as a "heuristic" theory by its author (Holland, 1966), it predicts that a person's behavior can be explained by the interaction of his personality pattern and his environment and that persons of a particular type seek vocational environments of that type. But as Holland maintains (1966, p. 90), "At this point the usefulness of the theory is only partially explored." And its interactional aspects in an institutional setting have been even less explored. From a methodological standpoint it may even be questionable whether personological and environmental scores with the same descriptive

label can shed much light on the psychological nuances of person-environment interactions. At this point more evidence is needed.

The theoretical dilemmas posed by interaction questions is well exemplified in an article by Werts and Watley (1969). These authors contrast the "environmental press" theory with the "relative deprivation" theory. The former theory predicts that strong environmental press for achievement in highly selective colleges will tend to increase student motivation for advanced training. The latter argues that students use their grades as an ability gauge, and if lower grades result from enrollment in a selective highly competitive college, this press will ultimately result in lowered aspirations for advanced training. The mediating variable of the ability self-concept thus assumes an important role. The authors provide a model for testing the contrasting predictions and present some preliminary evidence favoring relative deprivation theory. Astin (1969) responded to their contribution by pointing out some limitations in their data but acknowledging the importance of the methodological and substantive issues raised.

There were some studies that had little theoretical foundation of much consequence but which introduced some interesting methodological innovations to assess person-environment interactions. Lauterbach and Vielhaber (1966) asked a group of entering West Point cadets to complete the CCI in terms of West Point as they saw it and another entering group as they preferred it to be. The first instructions were designed to reflect expectations and the second to reflect needs. Congruency scores were computed between these scores and CCI press scores that were obtained from a group of experienced cadets who completed the form according to conventional instructions. Expectation-press congruence scores were significantly related to academic grades and so were need-press scores. However, need-press congruence scores were negatively related to academic grades, the opposite of what was predicted. Pervin (1967a, b) reported on two studies employing TAPE (Transactional Analysis of Personality and Environment), a semantic differential approach involving subject ratings of the following concepts on 52 scales: My College, My Self, Students, Faculty, Administration, Ideal College. Discrepancy scores representing the absolute difference in ratings of pairs of concepts were found to be correlated significantly with dissatisfaction with college (Pervin, 1967b). Distance scores (Osgood, Suci, and Tannenbaum, 1957) were similarly related (Pervin, 1967a). The method looks promising for the future, but it is too early to predict its ultimate contribution. The technique, of course, assesses "beta" press, and it is thus subject to all the advantages and disadvantages associated with that construct.

In terms of the five earlier suggested criteria for assessing research methodology directed to the study of person-environment interactions, the present area of endeavor could be evaluated as follows: It often con-



ceptualizes its research undertakings in terms of person-environment interactions, but just as often fails to implement its intent in its methods. It provides the tools for the measurement of situational differences, but these tools seem to be measuring the wrong variables. Its methods for describing person-environment interactions are not always logically or theoretically defensible and have little supporting evidence of their construct validity. It often succeeds, however, in providing a fairly comprehensive set of variables that adequately reflects the complexity of the environmental setting. And it would lend itself to many methods of multivariate statistical analysis if some of its other problems could be solved. Like the other two research areas reviewed, this one raises many difficult methodological issues. But it has developed rapidly in the scant 10-12 years that it has been in existence, and it is most worthy of the additional effort of researchers to insure the continuation of that progress in the future.

The three research areas in this paper have been reviewed to provide a glimpse of some selected current research undertakings that involve a person-environment interactional framework. The scientific and social relevance of this interactional frame of reference has been made most manifest as a result of contemporary events that have brought to our attention so forcefully the importance of situation and environmental context. The methodological problems of implementing any interactional approach are seen to be many and challenging, and progress may be slow. It may well be that psychological science as we know it will never prove methodologically adequate to the task in terms of any reasonable set of criteria. But prediction is the ultimate criterion of effectiveness in any science, and no one denies that predictive efficacy in psychology could be vastly improved. Perhaps improved interactional analyses will increase that predictive efficacy, and perhaps not. Whatever the final result, a move toward more frequent and more effective use of interactional approaches is surely a critical developmental task for the psychology of the future.

### Bibliography

- Amidon, Edmund and Flanders, Ned A. The Effects of Direct and Indirect Teacher Influence on Dependent-Prone Students Learning Geometry. *Journal of Educational Psychology* 52: 286-91; 1961.
- Amidon, Edmund J. and Flanders, Ned A. *The Role of the Teacher in the Classroom*. Minneapolis, Minn.: Paul S. Amidon and Assoc., 1963.
- Amidon, Edmund J. and Hough, John B. *Interaction Analysis: Theory, Research and Application*. Reading, Mass.: Addison-Wesley, 1967.
- Anderson, Harold H. and Brewer, Helen M. *Studies of Teachers' Classroom Personalities. I. Dominative and Socially Integrative Behavior of Kindergarten Teachers*. Applied Psychology Monographs No. 6. Stanford, Calif.: Stanford Univ. Press, 1945.
- Anderson, Harold H.; Brewer, Joseph E.; and Reed, Mary F. *Studies of Teachers' Classroom Personalities. III. Follow-Up Studies of the Effects of Dominative and Integrative Contacts on Children's Behavior*. Applied Psychology Monographs No. 11. Stanford, Calif.: Stanford Univ. Press, 1945.



- Astin, Alexander W. A Re-Examination of College Productivity. *Journal of Educational Psychology* 52: 173-78; 1961.
- Astin, Alexander W. An Empirical Characterization of Higher Educational Institutions. *Journal of Educational Psychology* 53: 224-35; 1962a.
- Astin, Alexander W. Influences on the Student's Motivation to Seek Advanced Training. Another Look. *Journal of Educational Psychology* 53: 303-309; 1962b.
- Astin, Alexander W. Differential College Effects on the Motivation of Talented Students to Obtain the Ph.D. *Journal of Educational Psychology* 54: 63-71; 1963a.
- Astin, Alexander W. Further Validation of the Environmental Assessment Technique. *Journal of Educational Psychology* 54: 217-26; 1963b.
- Astin, Alexander W. Classroom Environment in Different Fields of Study. *Journal of Educational Psychology* 56: 275-82; 1965a.
- Astin, Alexander W. Effects of Different College Environments on the Vocational Choices of High Aptitude Students. *Journal of Counseling Psychology* 12: 28-34; 1965b.
- Astin, Alexander W. Comment on A Student's Dilemma: Big Fish-Little Pond or Little Fish-Big Pond. *Journal of Counseling Psychology* 16: 20-22; 1969.
- Astin, Alexander W. and Holland, John L. The Environmental Assessment Technique: A Way to Measure College Environments. *Journal of Educational Psychology* 52: 308-16; 1961.
- Baker, Sheldon R. A Comparative Study of Perceptions of a University Environment between Honor and Nonhonor Freshmen Groups. *Educational and Psychological Measurement* 26: 973-76; 1966.
- Bracht, Glenn H. *The Relationship of Treatment Tasks, Personological Variables, and Dependent Variables to Aptitude—Treatment Interaction*. Doctor's thesis. Boulder: Univ. of Colo., 1969.
- Bracht, Glenn H. and Glass, Gene V. The External Validity of Experiments. *American Educational Research Journal* 5: 437-74; 1968.
- Campbell, James R. and Barnes, Cyrus W. Interaction Analysis—A Breakthrough. *Phi Delta Kappan* 50: 587-90; 1969.
- Campbell, Donald T. and Stanley, Julian C. Experimental and Quasi-Experimental Designs for Research on Teaching. *Handbook of Research on Teaching*. (Edited by Nate L. Gage.) Chicago: Rand McNally, 1963. Chapter 5, pp. 171-246.
- Carroll, John B. *Instructional Methods and Individual Differences. Learning and Individual Differences*. (Edited by Robert M. Gagné.) Columbus, Ohio: Charles E. Merrill, 1967. Pp. 40-44.
- Creager, John A. and Astin, Alexander W. Alternative Methods of Describing Characteristics of Colleges and Universities. *Educational and Psychological Measurement* 28: 719-34; 1968.
- Cronbach, Lee J. Processes Affecting Scores on "Understanding of Others" and "Assumed Similarity." *Psychological Bulletin* 52: 177-94; 1955.
- Cronbach, Lee J. The Two Disciplines of Scientific Psychology. *American Psychologist* 12: 671-84; 1957.
- Cronbach, Lee J. How Can Instruction Be Adapted to Individual Differences? *Learning and Individual Differences*. (Edited by Robert M. Gagné.) Columbus, Ohio: Charles E. Merrill, 1967. Chapter 2, pp. 23-39.
- Cronbach, Lee J. and Snow, Richard E. *Final Report: Individual Differences in Learning Ability as a Function of Instructional Variables*. Stanford, Calif.: School of Education, Stanford Univ., Mar. 1969.
- Di Vesta, Francis J. *Multivariable Study of Learning and Other Psychological Processes Related to Instruction*. Research proposal submitted to the Advanced Research Projects Agency, 1969.
- Flanders, Ned A. Teacher Behavior and In-Service Programs. *Educational Leadership* 21: 25-29; 1963.
- Flanders, Ned A. Some Relationships among Teacher Influence, Pupil Attitudes, and Achievement. *Contemporary Research on Teaching Effectiveness*. (Edited by Bruce J. Biddle and William J. Ellena.) New York: Holt, Rinehart and Winston, 1964. Chapter 7, pp. 196-231.
- Flanders, Ned A. *Teacher Influence, Pupil Attitudes, and Achievement*. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research

- Monograph No. 12. Washington, D.C.: Superintendent of Documents, Government Printing Office, 1965.
- Fromm, Erich. *Escape from Freedom*. New York: Rinehart, 1941.
- Fromm, Erich. *The Sane Society*. New York: Rinehart, 1955.
- Gage, N. L. Paradigms for Research on Teaching. *Handbook of Research on Teaching*. (Edited by N. L. Gage.) Chicago: Rand McNally, 1963. Chapter 3, pp. 99-141.
- Getzels, Jacob W. and Thelen, Herbert A. The Classroom Group as a Unique Social System. *The Dynamics of Instructional Groups*. (Edited by Nelson B. Henry.) Fifty-ninth Yearbook of the National Society for the Study of Education, Part II. Chicago: Univ. of Chicago Press, 1960. Chapter 4, pp. 53-82.
- Goldstein, Kurt. *The Organism*. New York: American Book Co., 1939.
- Herr, Edwin L. Differential Perceptions of Environmental Press by High School Students. *Personnel and Guidance Journal* 43: 678-86; 1965.
- Holland, John L. Undergraduate Origins of American Scientists. *Science* 126: 433-37; 1957.
- Holland, John L. *The Psychology of Vocational Choice: A Theory of Personality Types and Model Environments*. Waltham, Mass.: Blaisdell, 1966.
- Inkeles, Alex. Sociology and Psychology. *Psychology: A Study of a Science*. (Edited by Sigmund Koch.) New York: McGraw-Hill, 1963. Vol. 6, pp. 317-87.
- Jensen, Arthur R. Reinforcement Psychology and Individual Differences. *California Journal of Educational Research* 4: 174-78; 1962.
- Jensen, Arthur R. Varieties of Individual Differences in Learning. *Learning and Individual Differences*. (Edited by Robert M. Gagné.) Columbus, Ohio: Merrill, 1967. Chapter 6, pp. 117-35.
- Jensen, Arthur R. Social Class, Race and Genetics: Implications for Education. *American Educational Research Journal* 5: 1-42; 1968.
- Jensen, Arthur R. How Much Can We Boost IQ and Scholastic Achievement? *Harvard Educational Review* 39: 1-123; 1969.
- Kelly, George A. *The Psychology of Personal Constructs*. New York: Norton, 1955.
- Knapp, R. H. and Goodrich, H. B. *Origins of American Scientists*. Chicago: Univ. of Chicago Press, 1952.
- Knapp, R. H. and Greenbaum, J. J. *The Younger American Scholar: His Collegiate Origins*. Chicago: Univ. of Chicago Press, 1953.
- Lauterbach, Carl G. and Vielhaber, David P. Need-Press and Expectation-Press Indices as Predictors of College Achievement. *Educational and Psychological Measurement* 26: 965-72; 1966.
- Lewin, Kurt. *A Dynamic Theory of Personality*. New York: McGraw, 1935.
- Lewin, Kurt. *Principles of Topological Psychology*. New York: McGraw, 1936.
- Marks, Edmond. Personality and Motivational Factors in Responses to an Environmental Description Scale. *Journal of Educational Psychology* 59: 267-74; 1968.
- Maslow, Abraham H. Some Educational Implications of the Humanistic Psychologies. *Harvard Educational Review* 38: 685-96; 1968.
- McFee, Anne. The Relation of Students' Needs to their Perception of a College Environment. *Journal of Educational Psychology* 52: 25-29; 1961.
- Medley, Donald M. and Mitzel, Harold E. A Technique for Measuring Classroom Behavior. *Journal of Educational Psychology* 49: 86-92; 1958.
- Medley, Donald M. and Mitzel, Harold E. Measuring Classroom Behavior by Systematic Observation. *Handbook of Research on Teaching*. (Edited by Nate L. Gage.) Chicago: Rand McNally, 1963. Chapter 6, pp. 247-328.
- Meux, Milton O. Studies of Learning in the School Setting. *Review of Educational Research* 37: 539-62; 1967.
- Michael, William B. and Boyer, Ernest L. Campus Environment. *Review of Educational Research* 35: 264-76; 1965.
- Mitchell, James V., Jr. The Identification of Student Personality Characteristics Related to Perceptions of the School Environment. *The School Review* 76: 50-59; 1968.
- Murray, Henry A. *Explorations in Personality*. New York: Oxford Univ. Press, 1938.
- Murray, Henry A. and Kluckhohn, Clyde. Outline of a Conception of Personality. *Personality in Nature, Society and Culture*. (Edited by Clyde Kluckhohn, Henry A. Murray, and David M. Schneider.) New York: Knapp, 1953. Chapter I.
- Nunally, Jum C.; Thistlethwaite, Donald L.; and Wolfe, Sharon. Factored Scales for

- Measuring Characteristics of College Environments. *Educational and Psychological Measurement* 23: 239-48; 1963.
- Osgood, Charles E.; Suci, George J.; and Tannenbaum, Percy H. *The Measurement of Meaning*. Urbana: Univ. of Ill. Press, 1957.
- Pace, C. Robert. Five College Environments. *College Board Review* 41: 24-28; 1960.
- Pace, C. Robert. *CUES: College and University Environment Scales*. Princeton, N. J.: Educational Testing Service, 1963.
- Pace, C. Robert and Stern, George G. An Approach to the Measurement of Psychological Characteristics of College Environments. *Journal of Educational Psychology* 49: 269-77; 1958.
- Panos, Robert J. and Astin, Alexander W. Attrition among College Students. *American Educational Research Journal* 5: 57-72; 1968.
- Pervin, Lawrence A. Satisfaction and Perceived Self-Environment Similarity: A Semantic Differential Study of Student-College Interaction. *Journal of Personality* 35: 623-34; 1967a.
- Pervin, Lawrence A. A Twenty-College Study of Student  $\times$  College Interaction Using TAPE. *Journal of Educational Psychology* 58: 290-302; 1967b.
- Pervin, Lawrence A. Performance and Satisfaction as a Function of Individual-Environment Fit. *Psychological Bulletin* 69: 56-58; 1968.
- Piaget, Jean. *The Origins of Intelligence in Children*. New York: International University Press, 1952.
- Rotter, Julian B. The Role of the Psychological Situation in Determining the Direction of Human Behavior. *Nebraska Symposium on Motivation—1955*. (Edited by Marshall Jones.) Lincoln, Neb.: Univ. of Neb. Press, 1955.
- Scott, W. A. Reliability of Content Analysis: the Case of Nominal Coding. *Public Opinion Quarterly* 19: 321-25; 1955.
- Sells, Saul B. An Interactionist Looks at the Environment. *American Psychologist* 18: 696-702; 1963.
- Spearman, Charles. *The Abilities of Man*. New York: Macmillan, 1927.
- Stern, George G. Congruence and Dissonance in the Ecology of College Students. *Student Medicine* 8: 304-39; 1960a.
- Stern, George G. Student Values and Their Relationship to the College Environment. *Research on College Students*. (Edited by H. T. Sprague.) Boulder, Colo.: Western Interstate Commission for Higher Education, 1960b. Pp. 67-104.
- Stern, George G. Environments for Learning. *The American College*. (Edited by Nevitt Sanford.) New York: Wiley, 1962a. Chapter 21, pp. 690-730.
- Stern, George G. The Measurement of Psychological Characteristics of Students and Learning Environments. *Measurement in Personality and Cognition*. (Edited by Samuel Messick and John Ross.) New York: Wiley, 1962b. Pp. 27-68.
- Stern, George G. Characteristics of the Intellectual Climate in College Environments. *Harvard Educational Review* 33: 5-41; 1963a.
- Stern, George G. *Scoring Instructions and College Norms: Activities Index, College Characteristic Index*. Syracuse, N.Y.: Psychological Research Center, 1963b.
- Stern, George G.; Stein, Morris I.; and Bloom, Benjamin. *Methods in Personality Assessment*. Glencoe, Ill.: The Free Press, 1956.
- Stricker, George. Interrelationships of Activities Index and College Characteristics Index Scores. *Journal of Counseling Psychology* 14: 368-70; 1967.
- Sullivan, Harry S. *The Interpersonal Theory of Psychiatry*. New York: Norton, 1953.
- Sullivan, Harry S. *Clinical Studies in Psychiatry*. New York: Norton, 1956.
- Thistlethwaite, Donald L. College Press and Student Achievement. *Journal of Educational Psychology* 50: 183-91; 1959.
- Thistlethwaite, Donald L. Fields of Study and the Development of Motivation to Seek Advanced Training. *Journal of Educational Psychology* 53: 53-64; 1962a.
- Thistlethwaite, Donald L. Rival Hypotheses for Explaining the Effects of Different Learning Environments. *Journal of Educational Psychology* 53: 310-15; 1962b.
- Thistlethwaite, Donald L. and Wheeler, Norman. Effects of Teacher and Peer Subcultures upon Student Aspirations. *Journal of Educational Psychology* 57: 35-47; 1966.
- Walker, Helen M. Methods of Research. *Review of Educational Research* 26: 323-44; 1956.

- Werts, Charles E. *The Study of College Environments Using Path Analysis*. Research Report 3. Evanston, Ill.: National Merit Scholarship Corporation, 1967.
- Werts, Charles E. The Partitioning of Variance in School Effects Studies. *American Educational Research Journal* 5: 311-18; 1968.
- Werts, Charles E. and Watley, Donovan J. Analyzing College Effects: Correlation vs. Regression. *American Educational Research Journal* 5: 585-98; 1968.
- Werts, Charles E. and Watley, Donovan J. A Student's Dilemma: Big Fish-Little Pond or Little Fish-Big Pond. *Journal of Counseling Psychology* 16: 14-19; 1969.
- Withall, John. Development of a Technique for the Measurement of Socio-Emotional Climate in Classrooms. *Journal of Experimental Education* 17: 347-61; 1949.
- Withall, John and Lewis, W. W. Social Interaction in the Classroom. *Handbook of Research on Teaching*. (Edited by Nate L. Gage.) Chicago: Rand McNally, 1963. Chapter 13, pp. 683-714.

## AUTHOR

JAMES V. MITCHELL, JR. Address: University of Rochester Title: Associate Dean for Graduate Studies Age: 44 Degrees: A.B., A.M., Ph.D., Univ. of Chicago Specialization: Educational Psychology and Research; Personality, Measurement, and School Learning Environments.



American Educational Research Association  
Washington, D. C.

# New and Forthcoming Publications from AERA

*Encyclopedia of Educational Research*, Fourth Edition. Robert Ebel, editor.

AERA's classic in the field. A new edition citing references for and describing research findings on topics in education from academic freedom to technical education. May, 1969. Single volume: 1,500 pages. Order directly from Macmillan Company, New York City. Price: \$30.

*AERA Curriculum Evaluation Monograph Series*. Robert E. Stake, editor.

- \*1. *Perspectives of Curriculum Evaluation*. Ralph Tyler, Robert Gagné, Michael Scriven. 1968; Now in second printing.
- \*2. *Evaluation Activities of Curriculum Projects: A Starting Point*. Hulda Grobman. Fall 1968.
- \*3. *Instructional Objectives*. W. James Popham, Elliot Eisner, Howard Sullivan, Louise Tyler. Fall 1969.
4. *Research Strategies for Evaluating Training*. Phillip H. DuBois and E. Douglas Mayer, editors. Spring 1970.
5. *Evaluation as a Tool in Curriculum Development: The IPI Evaluation Program*. C. M. Lindvall and Richard C. Cox. Spring 1970.

\*Available now.

Order all volumes directly from Rand-McNally and Company,  
P.O. Box 7600, Skokie, Illinois. Price: \$3.50; \$2.00 (paper).

AERA membership discount on all publications: 30%

This Chapter represents a departure in two ways from past statistical reviews published in the *Review of Educational Research*. First, rather than presenting a review of the literature, the authors were encouraged to evaluate current research and to give their own viewpoints as to the past and future practice of statistics. Second, it was apparent that no one person could accomplish this task because of the specialized nature of the field of Statistics today.

The chapter is divided into four sections. The topics presented are Bayesian statistics, multivariate statistics, non-parametric statistics and Monte Carlo methods. Each section was independently authored with no effort made towards integration. The authors do not claim comprehensive coverage of even the areas of statistics represented here.

## I. BAYESIAN STATISTICS

DONALD L. MEYER

*Syracuse University*

A detailed philosophical discussion of Bayesian statistics is not given in this chapter, nor is an extensive list of references furnished. Rather, an overview of the Bayesian "movement" will be represented together with some thoughts on future developments.

It is convenient to think of modern-day Bayesian statistics as developing in three slightly overlapping stages. The first stage occurred in the 1950's and can be characterized as a period of philosophical foundation building. The second stage occurred in the 1960's and was a period of extensive derivation of posterior distributions applicable to many of the problems familiar to users of traditional or "classical" statistics. The third stage is just beginning and hopefully will be characterized by a wealth of applications of the thought and theorems of the first two stages.

### *Philosophical Foundations*

Savage's (1954), *The Foundations of Statistics*, is the primary reference on the foundations of Bayesian inference. From basic axioms of choice and ordering of preferences, the logical structure of a theory of subjective probability was developed and, in particular, the existence of prior proba-

bility distributions was established. Of course, these thoughts did not wholly originate with Savage, but were the combined effort of writers going back hundreds of years to Bayes himself. Kyberg and Smokler (1964) collected several papers and an extensive bibliography on subjective probability spanning seventy years. However, most readers would probably learn more from the modern treatment given the subject by Savage (1954), Good (1950, 1965) or Jeffreys (1961). The latter two authors provide direct applications of the theory to traditional problems of goodness of fit, tests of means, etc.

Contemporary critics of Bayesian methods do not attack the logic or the basic axioms of the theory. Rather, they seem to object to the "over-formalism" of procedures and to the fact that the theory is subjective. The latter finds expression in the task of specifying a prior distribution for the parameters under investigation. This specification is both difficult and troublesome, but it is unavoidable. The Neyman-Pearson theory was developed to circumvent the problem of procedures which, by and large, have certain optimal characteristics measured over the long run. Unfortunately, the results of any particular analysis cannot be interpreted sensibly within the theory. For example, a confidence interval on a mean of a normal population (with known variance) is calculated as  $20 < \mu < 30$ , say, with a confidence coefficient of .95. Many of us have been taught that this .95 is not a probability which can be applied to this particular interval; that  $\mu$  is either between 20 and 30 or it is not. Of what use, then, is this interval? Most of us believe that data, wisely collected, tells something about the problem under investigation so that we would respond that we believe that  $\mu$  is between 20 and 30. How strong our belief should be is an open question. One of the purposes of statistical theory is to guide one in making statements with known amounts of risk, but the Neyman-Pearson theory does not permit us to state the risk. Because of the long-run property that 95 per cent of intervals calculated in the same way would contain  $\mu$ , one is tempted to say that the risk or probability of being wrong given the data is .05. That is, one would be indifferent between a bet that one particular playing card could be picked out of twenty dealt haphazardly face down and a bet that  $\mu$  is not between 20 and 30.

Suppose a researcher did say this, forgetting Neyman-Pearson for the moment. Further, suppose he were asked, before seeing the present data, his preference for a bet that  $\mu$  is between 20 and 30 or that  $\mu$  is between 10 and 20. It can be shown that to be consistent he should respond that he would be indifferent between these two bets. That is, if a person samples from a normal population and takes the resulting .95 confidence interval for  $\mu$  as an interval he would be willing to give 19 to 1 odds on, then he should have been willing to give even odds that  $\mu$  is in any two equal length intervals before sampling.

In Bayesian terminology, as the prior distribution for  $\mu$  approaches a

uniform distribution, then the .95 posterior probability interval for  $\mu$  approaches the .95 confidence interval. That is, a prior distribution for  $\mu$  carries an implication about the data which will be observed, and the distribution of confidence intervals. The way one bets after seeing the data implies what one would have bet *a priori*.

Nobody says how one should bet on the .95 confidence interval, but what Bayesian gives is an "if-then" relationship. If one is only willing to give even odds on the .95 interval, then this carries a different implication from that given above. It is exactly this freedom of choice which is objected to by the critics. Why an elucidation of an "if-then" relationship should be objected to is difficult to understand.

### Posterior Distributions

Many, if not most, of the traditional techniques and statistical methods have been investigated using Bayesian theory. This investigation has taken two forms. First, the particular prior distributions which lead to posterior distributions such that the  $(1-\alpha)$  probability intervals are equivalent to the  $(1-\alpha)$  confidence intervals have been derived. Second, a class of prior and posterior distributions called natural conjugate Bayes densities (NCBD) unique to a model and method of sampling have been catalogued. An NCBD has the property that if the prior distribution is a member of the class, then the posterior distribution is another member of the class.

As an example of the first type, consider the case of sampling  $n$  observations from a normal distribution with unknown variance. The usual confidence interval for  $\mu$  utilizes the  $t$ -distribution with  $n-1$  degrees of freedom and the confidence interval for  $\sigma^2$  is based on the chi-square distribution with  $n-1$  degrees of freedom. What prior distribution would lead to posterior distributions such that the probability intervals would equal these confidence intervals? The answer is that the joint prior distribution for  $\mu$  and  $\sigma^2$  is such that  $\mu$  is independent of  $\sigma^2$  and that  $\mu$  and  $\ln\sigma^2$  are each distributed uniformly. This has been referred to as the "indifference" distribution characterizing ignorance. Note that this prior is somewhat restrictive in not allowing correlation between  $\mu$  and  $\sigma^2$ , and it is not at all clear why "indifference" should refer to  $\ln\sigma^2$  and not  $\sigma^2$ . However, so be it if one interprets the confidence intervals in the way discussed previously.

The NCBD for this model is the normal-gamma distribution. This is a bivariate distribution such that  $\sigma^2$  is distributed in the gamma form with  $\mu$  distributed normally with variance depending on  $\sigma^2$ . After sampling, the posterior distribution is another member of the normal-gamma family. One of the parameters of the distribution can be conceptualized as a prior sample size,  $n'$ . The corresponding parameter in the posterior distribution is  $n''$  where  $n'' = n' + n$  and  $n$  is the present sample size. Therefore, picking a prior is, essentially, choosing a starting point for the accumulation of knowl-



edge. If  $n' = 0$ , the indifference distribution noted above results. A remarkable consistency!

A recent conference on industrial applications of Bayesian methods brought out the fact that some engineers after seeing a certain set of data for a particular problem would not bet according to the applicable NCBD. The speaker was troubled about this and implied that this was a limitation of Bayesian methods. It is clear, however, that either these engineers did not change their belief the way they should or that their prior distribution was not one of the NCBD class. Further investigation and experience with this problem should be rewarding rather than limiting.

Relevant references for indifference and NCBD distributions can be found in the review by Meyer (1966). See especially Raiffa and Schlaiffer (1961) which is now in paperback form.

No real surprises have been found in the sense that traditional methods are equivalent to Bayesian methods if, generally speaking, the same type of indifference distributions previously referred to are used as prior distributions. What has resulted from Bayesian research, however, is a widening of available procedures and, more important, a different way of thinking.

For example, since all knowledge for a particular problem is summarized by the posterior distribution, one has a method of *describing* the model independent of the necessity to consider a decision process and the associated loss function. Traditional methods with heavy emphasis on the testing of hypotheses confuse descriptions with decisions. Some argue that a significance level can be used as an indication of the weight of evidence against an hypothesis. The use of a test of significance in this way has been suggested by Jerome Cornfield in his Phi Delta Kappa symposium paper to rest on the unstated postulate that "All hypotheses rejected at the same critical level have equal amounts of evidence against them." Cornfield gives a counter-example and discusses the Bayesian solution. See his paper in Meyer and Collier (1970).

With Bayesian methods, if one is interested in making a decision then the loss function together with the posterior distribution is considered, but one can stop at the description stage. This is in the spirit of the paper by Tukey (1960) who argued for a theory of "conclusions" rather than decisions.

A Bayesian also views estimation problems differently. Traditional methods must be concerned with long-run properties of statistical estimators over samples not observed. This does not arise in Bayesian methodology since only the actual observations are used. Put another way, integration is over the parameter set, not the sample space. The test of various properties of summary statistics is always with reference to the posterior distribution. A sufficient statistic, for example, is one which results in the same posterior distribution as the original observations.

A good example of the Bayesian way of viewing old problems was given by Hill (1967). He considered the problem of inference given a negative variance estimate in the one-factor components of variance model. If a negative estimate occurs, then belief in the usual model is weakened. Hill showed that an alternative model which can gain belief is one which allows for negatively correlated errors. In other words our distribution of belief concerning the model, itself, is changed. Previous writers on this problem suggested that the model may be incorrect, but Hill has shown explicitly not only how this can happen but also the inter-relationships of the two models.

See Pitz (1968) for an example of a Bayesian view of a hypothesis test applied to a psychological experiment.

### *The Future*

Very little use of Bayesian methods in the behavioral sciences is currently evident. Wider application can be expected when Bayesian methods become available in more usable form. One of the practical difficulties, for example, is the need for tables of relatively unfamiliar distributions. Publication of textbooks will help. New books with direct Bayesian applications have been written by Sasaki (1968), Savage (1968) and Schmitt (1969).

One type of application of Bayesian methods which is quite exciting is to the design of experiments and sample surveys. During the planning stage of an investigation, one must make decisions about method, allocation of resources, etc., using prior knowledge and expectations. Everyone, regardless of persuasion, must agree with this. A Bayesian can attack this problem by considering a prior distribution for both primary parameters and nuisance parameters such as non-response rates, measurement error and the like. The choice of design would be the one which had the greatest expectation of affecting the experimenter's beliefs about the values of the primary parameters. One criterion for choosing a design might be to choose the design which has the smallest expected variance of the posterior distribution of the parameter under study.

After the study is performed, one is able to report his results with both the uncertainty and knowledge concerning the nuisance parameters taken into account. That is, if measurement error is thought to have a positive biasing effect, then the final probability interval would be adjusted in the negative direction with probably wider limits as a result of the additional uncertainty. By formalizing the process of design, the subjective decisions necessary are made explicit; study of the inter-relationships of these decisions should be rewarding. See Glass and Maguire (1968) who applied Bayesian methods to analyzing time-series quasi-experiments and Brown (1967) for an application to sample surveys.

A note of caution is in order. The teaching of Bayesian statistics is a good deal more difficult than teaching traditional methods. The students need to know more probability theory and more distribution theory. Since Bayesian theory is more resistant to a cookbook approach, a deeper understanding of underlying processes is necessary. This may be an advantage since it may reduce the common misuse of genuinely difficult concepts by those who have been taught to summarize data by reading the .05 value of  $t$ .

## II. NONPARAMETRIC METHODS

MARYELLEN McSWEENEY

*Michigan State University*

LEONARD A. MARASCUILO

*University of California*

### *Robustness, Efficiency and "Assumption-Freer" Statistics*

In his review of Bayesian statistics, Meyer (1966) suggested that an evolution of statistical theory and practice would remove the qualifiers *Fisherian*, *Pearsonian*, *classical*, and *Bayesian* from statistics. The same assertion can be made about the qualifiers *parametric* and *nonparametric*. Noether (1967) and Hoeffding (1968) noted that the parametric-nonparametric distinction is neither precise nor informative. Most tests that assume normality or homogeneity of variance are asymptotically nonparametric for very large samples. Their nominal significance levels are approximately correct and their power approaches unity although the samples are drawn from populations which violate parametric assumptions; moreover, the statistical power of the nonparametric tests depends on the characteristics of the distributions sampled so that these tests are not genuinely "distribution-free."

The blurring of the parametric-nonparametric distinction is not a commentary on the sterility of the "assumption-freer" approach, but a consequence of greater theoretical interest in the behavior of inferential techniques under conditions of nonnormality and heterogeneity of variance. (The term "assumption-freer" statistics was introduced by Hans Ury (1967) in response to Noether's letter, "Needed—A New Name.") When the term "non-parametric" was introduced by Wolfowitz (1942), only a few simple rank tests were available. Although several of these early tests had been developed by social scientists (e.g., Friedman, 1937; Festinger, 1946), nonparametric techniques did not commonly appear in methodological articles and statistical texts directed toward the social scientists until the early 1950's. Blum and Fattu (1954) ascribe the rela-

tively infrequent use of nonparametric statistics in educational research to the absence of textbook coverage of nonparametric statistics at the intermediate level prior to 1953. The attention given to nonparametric statistics in various review articles (Johnson and Moonan, 1951; Moses, 1952; Blum and Fattu, 1954) and the extensive textbook treatment of nonparametric techniques by Festinger and Katz (1953) and Walker and Lev (1953) were largely responsible for introducing these techniques to a wide audience of potential users in education. Although the early reviews of nonparametric techniques stressed their "assumption-freer" properties rather than their ease of application, the techniques soon gained favor among researchers as "quick and dirty" procedures (Tate and Clelland, 1957). Since practitioners emphasized the tentative nature of the nonparametric tests, they gave little attention to the relative efficiency of the nonparametric procedures when compared with their parametric analogues. More recently, renewed interest in the robustness of parametric tests and more extensive study of the properties of rank statistics and other order statistics have prompted statisticians to compare the efficiency of nonparametric procedures relative to their parametric analogues.

Hodges and Lehmann (1956) established a lower bound for the asymptotic relative efficiency of the Wilcoxon test to the two-sample  $t$ -test when both tests are used to detect differences in the location of two populations. Regardless of the populations sampled, the Wilcoxon test always has an asymptotic efficiency of at least 0.864 relative to the  $t$ -test. The numerical value for the efficiency of test  $A$  relative to test  $B$  may be loosely interpreted as the inverse ratio of the sample sizes ( $n_B / n_A$ ) required for the two tests to have the same power against the same alternative at the same  $\alpha$  level. Thus, 0.864 implies that a  $t$ -test based on 86.4 observations would have the same power as a Wilcoxon test based on 100 observations. This lower bound applies to the atypical distribution,

$$f(x) = 3\sqrt{5(5-x^2)} / 20, x^2 \leq 5.$$

the relative efficiency of the Wilcoxon test to the  $t$ -test is actually much higher than the 0.864 minimum for commonly encountered distributions (e.g., the asymptotic relative efficiency is 0.955 for normal distributions, 1.000 for uniform, and 1.047 for logistic distributions). Chernoff and Savage (1958) established a lower bound of 1.000 for the efficiency of the two-sample normal scores test to the  $t$ -test. The efficiencies cited imply that the Wilcoxon test is often as powerful as the  $t$ -test and never appreciably worse, while the normal scores test is *always at least as powerful* as the  $t$ -test.

These results are probably somewhat startling to the practitioner who has routinely selected a parametric test in preference to a nonparametric one because of an erroneous belief that the latter always discards informa-



tion and is therefore less powerful. Even more surprising is Klotz's (1963, p. 631) preference for the routine use of rank tests:

Because of the extremely high efficiency of the nonparametric tests relative to the  $t$  in the region of interest, it is the author's opinion that the nonparametric tests would be preferred to the  $t$  in almost all practical situations. The exactness of the null distribution, good power for a wide class of shift alternatives, and the negligible loss in efficiency on the home ground of the  $t$ -test support this conclusion.

A critic of Klotz's statement might reasonably wonder, "What about robustness? Aren't most parametric tests for location insensitive to non-normality and heterogeneity of variance?" A partial answer has been supplied by mathematical and empirical studies of the parametric and nonparametric tests for various nonnormal populations and dissimilar variances. Three particularly relevant studies of the robustness of parametric procedures are Pratt's (1964) study of various two-sample tests, Holloway and Dunn's (1967) study of Hotelling's  $T^2$ , and Baker and Collier's (1968) study of the  $F$ -test in a randomized block design.

Pratt compared the  $t$ -, median, Wilcoxon, and normal scores tests as tests of location in symmetric, unimodal distributions with unequal variances. Under these conditions, the  $t$ -test is less robust than the nonparametric tests if the sample sizes are unequal, but the actual significance levels of all of the tests are somewhat affected by inequality of variance. Holloway and Dunn reported that both the actual  $\alpha$  level and the power of  $T^2$  are adversely affected by inequality of the covariance matrices. Even equating the sample sizes for the two groups will not restore the power of the test if the covariance matrices are unequal. Baker and Collier found that the  $F$ -test for randomized blocks is also affected by inequality of variance in the blocks as well as by treatment-by-block interaction, but the test is insensitive to leptokurtosis.

### *Nonparametric Analysis of Variance: Univariate and Multivariate*

Nonparametric tests have frequently received only perfunctory attention from behavioral researchers who wish to analyze the data obtained in multi-factor experiments. Researchers who categorize nonparametric tests as "quick and dirty" are unlikely to seek a nonparametric means of analyzing a complex design. Moreover, those researchers who do look for a distribution-free method of analysis find that the familiar nonparametric analyses of variance are either inefficient or inadequate.

Extended median tests such as the Mood-Brown and Wilson tests are inefficient competitors of the parametric procedures since they retain only the sign and not the relative magnitude of the observations. The more efficient Kruskal-Wallis and Friedman tests use the relative magni-

tude of the observations to preserve power, but they compete with the parametric procedures only when the experimental designs are relatively simple. The researcher who plans a completely random multifactor study or who nests subjects in unequal size groups does not have a readily available analytic technique if he is unable to meet the assumptions for a parametric analysis.

This apparent void in the area of distribution-free ANOVA does not indicate that nonparametric techniques are impossible but that the theoretical problems are complex and that statisticians are lax in presenting new techniques to a wide audience of potential users. Some of the newer nonparametric techniques appear only in technical statistical literature; others are computationally so burdensome that they defy description as "quick and dirty." Recent publications in applied statistics journals and intermediate texts on rank order statistics (Noether, 1967; Bradley, 1968; Kraft and Van Eeden, 1968; Hájek, 1969) are attempts to fill this void by presenting new methods and, in some cases, making the appropriate computer programs available.

Among the nonparametric advances in the analysis of complex designs, three closely related approaches seem especially promising: the method of aligned ranks, nonparametric MANOVA, and planned contrasts. Hodges and Lehmann (1962) introduced the alignment and subsequent ranking of observations in the randomized block design to eliminate the confounding of treatment and block effects. The aligned rank procedure yields highly efficient tests of treatment effects for generalized randomized blocks, balanced incomplete blocks, and Latin squares. Alignment is appealing because it enables the researcher to increase the precision of his experiment by blocking, without restricting him to the use of equal-sized blocks. The use of aligned ranks first requires complete randomization within each block, and is followed by subtraction of the block mean or median from each observation in the block. Despite the freedom from distribution assumptions of this procedure, the asymptotic efficiency of test statistics based on the ranking of aligned observations or the replacement of these observations by normal scores is high when compared with classical ANOVA. Sen (1968) extended the aligned rank procedure to a two-way layout with symmetrically dependent errors and noted that high efficiency relative to classical ANOVA was maintained. Once again, the lower bound on efficiency for ranks was 0.864 and for normal scores, 1.000.

Sen's study of symmetrically dependent errors led to Koch and Sen's (1968) consideration of nonparametric ANOVA techniques for the mixed model with  $n$  randomly chosen subjects who respond once to each of  $p$  treatments. Univariate parametric analysis of the mixed model requires that the errors be multivariate normal with variances  $\sigma^2$  and covariances  $\rho\sigma^2$ . A parametric analysis using Hotelling's  $T^2$  requires 'additivity' of

the subject effects. Although  $T^2$  and the univariate  $F$ -tests are not especially robust against inhomogeneity of variance, the rank tests of Koch and Sen are insensitive to violation of the parametric assumptions. Koch and Sen present several variations of their basic statistic to accommodate different assumptions about the covariance structure of the treatments and the presence of subject-by-treatment interaction.

Koch (1969) extended this work to provide a nonparametric analysis for a repeated measures design with subjects nested within groups. Koch analyzed this frequently used design with  $p$  measures per subject and  $n_j$  subjects per group for each of  $J$  groups as a nonparametric one way MANOVA. The freedom of this procedure from common parametric assumptions, its versatility in handling groups of unequal size, and its high efficiency when compared with classical ANOVA and MANOVA make nonparametric MANOVA especially useful.

Planned contrasts are also used in distribution-free ANOVA. Such an approach identifies linear combinations of the population parameters (contrasts) and creates tests of the nullity of these contrasts as a substitute for omnibus hypothesis tests (Lehmann, 1963a, b, and c; Bradley, 1968, pp. 138-41; Kao, 1969). The procedure is noteworthy not for its novelty, but for its versatility. It enables the researcher to create distribution-free tests of specific subhypotheses and to combine these subtests if desired. If, in addition, the error rate for the family of subtests is controlled, the researcher has an effective technique for asking more specific questions of his data than those permitted by an omnibus hypothesis test.

Lehmann's (1963a, b, and c) application of planned contrasts to create distribution-free estimates of the treatment effects in ANOVA is a major theoretical contribution to nonparametric statistics, but is not yet a practical alternative for most researchers. Lehmann's procedure yields robust and efficient estimates for contrasts and hypothesis tests, but the computational labors required and the possible instability of the estimate of internal variability obtained by this method may give the behavioral researcher pause.

### *Nonparametric Estimation with a Controlled Error Rate*

In many experiments requiring a comparison of  $K$  treatments the researcher is not content with disproving equality of the treatments; he is typically interested in identifying differences among the treatments and in estimating the magnitude of the treatment effects. Until the late 1950's the user of nonparametric statistics was forced to rely on multiple two-sample tests to identify specific relationships among the treatments in a  $K$ -treatment design. Use of multiple Mann-Whitney two-sample tests after a significant Kruskal-Wallis test was a common practice which ignored the interdependence of the two-sample tests and the enlarged error rate for the family of comparisons.

The difficulties encountered with the multiple two-sample tests have been circumvented by nonparametric multiple comparison procedures which have their parametric analogues in the Tukey and Scheffé procedures and in the use of the Bonferroni inequality. Tables of confidence coefficients based on the exact sampling distributions of nonparametric tests have been developed by Nemenyi (1963), McDonald and Thompson (1967), and Tobach et al. (1967). Most of these exact tables are restricted to the paired-comparison of treatments, to relatively small sample sizes, and to simple experimental designs.

Other nonparametric multiple comparisons are based on large-sample distribution theory (Steel, 1960; Dunn, 1964; Gold, 1963; Goodman, 1965; Rosenthal and Ferguson, 1965; Marascuilo, 1966; and Marascuilo and McSweeney, 1967). The resulting confidence intervals yield *approximate* rather than *exact* probability statements. Such statements are typically conservative and the confidence intervals are longer than those obtained with the exact procedures (Miller, 1966). Despite the greater length of these intervals, the large-sample approximations permit the examination of simple and complex contrasts, employ readily available tables, and handle large sample sizes without complicating the analysis.

Some nonparametric procedures are analogous to the parametric procedures of Tukey and Scheffé in that the user can examine as many comparisons as he wishes with the assurance that the error rate for the entire family of comparisons will not exceed a predetermined value (Miller, 1966; Marascuilo, 1966). Other nonparametric multiple comparison procedures control the error rate by means of the Bonferroni inequality (Dunn, 1964; McDonald and Thompson, 1967; Jensen et al., 1968). This inequality states that for  $L$  comparisons the experiment-wise error rate ( $\alpha$ ) is no greater than the sum of the error rates per comparison ( $\alpha_j'$ ),

$$\alpha < \sum_{j=1}^L \alpha_j'.$$

When nonparametric techniques using the Bonferroni inequality are compared with the nonparametric procedures of the Tukey-Scheffé type, the former techniques are usually found to give shorter confidence intervals unless the number of comparisons made is very large. Offsetting this advantage of the Bonferroni techniques is the relative inaccessibility of the necessary tables of intermediate percentiles for exact distributions and large-sample approximations. Even when such tables are available for large-sample approximations, the researcher may hesitate to use the Bonferroni inequality when  $\alpha_j' < .01$  because the approximations are usually less accurate in the extreme tails of the sampling distributions than at the moderate significance levels.

Multiple comparisons based on large-sample theory and employing



an experiment-wise error rate seem to be the most versatile of the non-parametric multiple comparison procedures. Their insensitivity to non-normality and inequality of variance makes these procedures desirable competitors of the parametric procedures of Tukey and Scheffé. Their simplicity and their use of readily available tables make the asymptotic procedures more practical than the exact nonparametric procedures at present. Increased use of computers to generate exact tables and to study the quality of the large-sample approximations to the exact sampling distributions may shift the balance in favor of procedures based on the exact sampling distributions. Regardless of the future direction taken by nonparametric multiple comparisons, the mere fact that the researcher has several alternatives to the uncontrolled use of multiple two-sample tests is a signal advance in nonparametric statistics.

### *Conclusion*

Nonparametric statistics evolved from a limited set of simple rank analogues of parametric tests to a diverse collection of mathematically sophisticated techniques. In the course of this evolution, the properties that originally made nonparametric statistics popular, ease and speed of computation, have largely been lost. The newer nonparametric procedures, typified by the ANOVA and estimation techniques described in this section, are neither quick nor easy, but they are extremely versatile and highly efficient. No matter how powerful the newer nonparametric techniques are, their usefulness will depend ultimately on the diligence of applied statisticians in explaining them and on the receptivity of behavioral researchers to these methodological advances.

## III. MONTE CARLO METHODS

RAYMOND O. COLLIER, JR.

ROBERT C. LARSON

*University of Minnesota*

The basic references from which the structure of this section was formulated are the symposium reports on Monte Carlo methods edited by Forsythe and Householder (1951) and Meyer (1956); the definitive monograph by Hammersley and Handscomb (1964); and a recent book on simulation by Naylor, Balintfy, Burdick and Chu (1968). In fact, if one were to study these references, he would probably encounter most of the critical material and references on Monte Carlo methods and not be motivated to read further below.

## General Objectives of Monte Carlo Methods

An application of "Monte Carlo methods," a "Monte Carlo investigation," in its primitive form consists of empirically sampling from a *supply* distribution to obtain an empirical sampling distribution and certain characteristics of this distribution for a random variable of interest. Since the objective of the process is to produce empirical approximations of the sampling distribution and its characteristics, Monte Carlo methods may be viewed as a special simulation technique (Naylor et al., 1968). The basic motivation for using Monte Carlo methods is either to approximate the solution of a problem which is based on a probabilistic structure (an activity referred to by some as "model sampling" or simply as "empirical sampling") or to approximate the solution of a deterministic problem by noting that the solution corresponds to a characteristic of a probability distribution of a random variable. Most researchers would agree that if analytic ("exact") solutions are available, they should be used instead of Monte Carlo methods unless the amount of labor per degree of exactness is unacceptably great.

As an example of the first kind of problem, suppose it is desired to obtain the expected value of the sample mean  $\mu_{\bar{x}}$ , computed from samples of size  $N$  from a population with density function  $f(x)$ . One can obtain an empirical distribution of sample means based on  $n$  replications and take  $\hat{\mu}_{\bar{x}} = \Sigma \bar{x}/n$  as an estimator of the desired expectation,  $\mu_{\bar{x}}$ . Among the Monte Carlo results also available in this example are answers to such questions as what is an estimate of the probability that  $x$  is greater than a number  $k$ ?

However, in contrast, as an example of the second kind of problem, suppose that in a certain *deterministic* problem a solution,  $y$ , is produced such that  $y = \int xf(x)dx$ . Because  $y = \mu_{\bar{x}}$ , one can use the previous Monte Carlo result and take  $\hat{\mu}_{\bar{x}}$  as an approximation to (estimator of)  $y$ . This procedure would be of value if  $\int xf(x)dx$  were complicated and the integration of the product difficult. Since the great majority of Monte Carlo investigations encountered in applied statistics are made on problems of the first type, the remarks in this section will be directed to this sort of problem.

It is interesting to note that the "model sampling" endeavor of Monte Carlo methods is not at all new. Early statisticians used this technique in teaching students applied probability to bolster their belief in theory, and even to elicit theoretical unknowns. Examples of early model sampling investigations are provided by W. S. Gossett in 1908 in his investigation of the distribution of the correlation coefficient and by Eden and Yates (1933) in their investigation of the permutation distribution of a function of a variance ratio. However, the realization of the solution central to the second type of problem is not much older than 1944 and

along with the term "Monte Carlo" is due to Ulam, Von Neuman, Fermi and colleagues in their work on the Manhattan Project.

### *Basic Sub-Topics of Monte Carlo Investigations*

Kahn (1955) pointed out that the three main points of importance in any Monte Carlo study are: 1) choice or analogization of the desired probability basis, 2) generation of sample values on a given piece of machinery, almost always a computer (but see Hammersley and Morton, 1954), and 3) design and use of procedures to reduce the variance of the estimate of the characteristic of interest. A decision made on 1) is basic to a specification of the problem itself and can offer challenging hurdles in the use of empirical methods. Presumably an investigator should be cognizant of the basic probability process about which he desires answers. In actual practice, however, the help which he often needs to formulate basic probability structures might be obtained from a mathematician, a statistician or a subject-field researcher. The following references are helpful in specifying basic and general probability structures: Ledley (1965), Raiffa and Schlaifer (1961), and Abramowitz and Stegun (1955), but for more specificity one will want to consult other sources such as actual Monte Carlo analyses in his area of interest.

Kahn's subtopic 2), the generation of sample values (realizations of random variables) for use in simulating probability processes by computer or by hand, can be done by actually storing in some receptacle a set of observations with a particular distribution from which random samples are drawn or by constructing an algebraic function which produces "random" samples of observations from a desired distribution. In using either method, the generation of a set of random numbers and tests of their efficacy for the purpose *at hand* are of extreme importance.

Random numbers can be generated by resorting to a table of random numbers or by employing a random physical phenomenon or process, but Monte Carlo work done by computer usually employs a deterministic mathematical function in order to generate "psuedo-random" numbers rapidly which can be checked for various behaviors. Discussions of several recursive mathematical functions and the results of tests of randomness with some indications of their shortcomings can be found in Hull and Dobell (1962, 1964), Greenberger (1965), MacLaren and Marsaglia (1965) and Van Gelder (1967). Theoretically, a wide class of probability distributions can be obtained by transforming observations from a random number generator (i.e., a uniformly distributed random variable), but in practice a variety of different methods becomes useful. This is illustrated in the papers of Marsaglia (1961, 1963, 1966), Marsaglia and Bray (1964) and Marsaglia and MacLaren (1964), who discuss methods for generating random discrete, exponential, and normal variables as well as random

variables with arbitrary distributions. Many of these methods exist today in computer libraries.

Kahn's last subtopic 3), techniques of variance reduction—although critically important, is much too neglected in Monte Carlo investigations on applied statistical matters by behavioral researchers. Recall the previously given example of a Monte Carlo analysis on the expected value of sample means. Suppose one is interested in finding the expected value of the distribution of sample means for random samples of size 10 from a normal population with mean 0 and variance 1. A specification of the probability basis as required in subtopic 1) states that  $f(x)$  is  $N(0,1)$  and that a random sample  $x_1, x_2, \dots, x_{10}$  from this population has joint probability (density) equal to

$$\prod_{i=1}^{10} (2\pi)^{-1/2} e^{-x_i^2/2}$$

Attention to subtopic 2) would suggest using a random normal generator with mean 0 and variance 1 to generate 10 independently distributed  $x$ 's and compute  $\bar{x}$ , repeat this process  $n$  times and compute  $\hat{\mu}_{\bar{x}} = \Sigma \bar{x}/n$  as an estimator of  $\mu_{\bar{x}}$ . Now it is known that the variance of the estimate  $\hat{\mu}_{\bar{x}}$  is  $\sigma_{\bar{x}}^2 = \sigma_x^2/n = \sigma_x^2/10n$  and in this case is  $1/10n$ . It is clear that one can

reduce the variance of  $\hat{\mu}_{\bar{x}}$  by increasing  $n$ ; however, more sophisticated and more effective methods for reducing variance of estimates for a given number of replications exist (see Kahn, 1955; Trotter and Tukey, 1956; Hammersley and Handscomb, 1964). Those procedures include the use of importance sampling, Russian roulette and splitting, expected values, correlation and regression, systematic sampling, and stratified sampling. All of these techniques capitalize on certain specific characteristics of the problem which make it possible to reduce the variance of the estimate by optimal sampling or by biasing the problem, i.e., using a different distribution than the one suggested by the problem, so that the variance of the estimate is reduced while making it possible to correct for bias at a later stage or to use weighted functions of various sub-estimators to reduce the variance of the estimate.

Contrary to the suggestions of various writers on sophisticated Monte Carlo methods, most investigators in behavioral areas use only the cruder, more obvious techniques in Monte Carlo analyses. Perhaps, because of ready availability of computer time, it is often simpler to run crude analyses than construct and employ more complicated optimal variance reducing schemes.



### *Design of More Intricate Monte Carlo Analyses*

The planning of multi-factor Monte Carlo computer investigations can be facilitated by utilizing the concepts of the design of experiments as presented by Kempthorne (1952), Federer (1955), Winer (1962), and Cox (1957). A good account of some of the problems in the design of *simulation* experiments in which the basic process is a business or economic one rather than a statistical or theoretical one is presented in Naylor et al. (1968), Naylor and Burdick (1966), and Smith (1968). It seems that the problems in designing optimal multi-factorial Monte Carlo experiments include a multiple decision scheme in which levels of independent and dependent variables are chosen so as to optimize a particular measure of information per cost.

A slightly different example can be cited which is connected to our previous example. Suppose the problem is to obtain the expected value of the sample median when various samples sizes represent the  $a$  levels of a factor  $A$  and various population symmetry indices represent the  $b$  levels of factor  $B$ . A plan is needed which will produce an estimate of the expected  $a \times b$  response surface of medians with a desired level of precision. Thus, the number of and values of the levels of Factors  $A$  and  $B$  must be specified and the number of replications to be made must be set so as to maximize the precision of estimation of this response surface for a given cost. (See Box, 1954; Box and Youle, 1955; and Box and Hunter, 1957. An important standard references in the sampling survey methodology literature is Hansen, Hurwitz and Madow, 1953.)

### *Some Examples of Monte Carlo Investigations on Particular Topics in Statistics*

To present some of the real flavor of Monte Carlo analysis, a few references will be cited in which this technique was used to obtain results on essentially statistical problems. Collier, Baker and Mandeville (1967) and Baker and Collier (1968) studied size and power of the  $F$ -test under permutation by means of Monte Carlo methods in an effort to demonstrate the robustness of the test under minimal assumptions. In addition, Collier, Baker, Mandeville and Hayes (1967) studied the size of standard  $F$ -tests and  $F$ -tests modified along the lines suggested by Greenhouse and Geisser (1959) in certain repeated measures designs under several different variance-covariance structures and for various sample sizes. The Monte Carlo method was particularly useful in estimating the size of an  $\epsilon$ -test for which the nominal  $\alpha$ -rejection region was  $F > F[\nu_1\hat{\epsilon}, \nu_2\hat{\epsilon}]$  in which the degrees of freedom  $\nu_1\hat{\epsilon}$  and  $\nu_2\hat{\epsilon}$  were *random variables* depending on sample variances and covariances.

Recently Monte Carlo methods were used by Klahr (1969) to study the random sampling distribution of the "stress statistic" in Kruskal's

nonmetric multi-dimensional scaling procedure. Feldt (1969) derived an approximate test of equality of two Kuder-Richardson formula 20 coefficients and, as a validity check, obtained Monte Carlo results of the sampling distribution for the statistic used in this test. An empirical sampling investigation was made by Hudson and Krutchkoff (1968) on tests employing Satterthwaite's synthetic mean squares with degrees of freedom estimated from the data. In the area of sample survey methodology, Gates (1969) compared several estimators for the line transect sampling method by empirical means, and Rao and Bayless (1969) ran an empirical study on the stability of estimators of population totals and their variance estimators in an equal probability sampling system. Finally, an extensive collection of references has been made by Elvebach and Varma (1965) for the general area of *simulation* including Monte Carlo methods and specifically bioassay, biological systems, epidemic theory, fertility, followup and genetics, hospital and blood banks, medical diagnosis, psychology and psychiatry, and social science.

### *Concluding Remarks*

Monte Carlo methods have been found to be valuable devices by which estimates of characteristics of probability distributions can be made. Mature suggestions concerning the application of these procedures come from Hammersley and Handscomb (1964, p. 9).

In the last few years Monte Carlo methods have come back into favor. This is mainly due to better recognition of those problems in which it is the best, and sometimes the only available technique. Such problems have grown in number, partly because improved variance-reducing techniques recently discovered have made Monte Carlo efficient where it had previously been inefficient, and partly because Monte Carlo methods tend to flourish on problems that involve a mass of practical complications of the sort encountered more and more frequently as applied mathematics and operational research come to grips with actualities.

## IV. MULTIVARIATE ANALYSIS

MAURICE M. TATSUOKA

*University of Illinois*

During the past decade, there has been a sharp increase in the use of multivariate statistical techniques such as discriminant analysis and multivariate analysis of variance in educational research. Factor analysis has

long been a favorite tool in educational and psychological research. This trend is partially due to the effects of (a) the publicizing efforts of numerous authors, including Bock (1966), Bock and Haggard (1968), L. V. Jones (1966), Rulon (1951), Rulon and Brooks (1968), and Tatsuoka and Tiedeman (1954, 1963); (b) the increased availability of electronic computers; and (c) the publication of several excellent books devoted to computer programs, such as those by Cooley and Lohnes (1962) and Ralston and Wilf (1960, 1967).

Over and above these impetuses, however, there is the fact that educational research is inherently multivariate in nature. Educational outcomes are seldom measured against a single criterion variable. The alternatives open to the educational researcher are, thus, either to conduct a series of univariate analyses (one for each criterion variable) or to carry out a multivariate analysis incorporating all the criterion variables simultaneously. Seen in this light, the increased use of multivariate analysis in educational research, phenomenal though it is, still falls far short of its ideal rate, namely 99 + percent. The basis for this seemingly bombastic assertion is that the first alternative mentioned above is really not a valid one. A detailed explanation of why this is so would require more space than we can afford here; suffice it to say that, for one thing, the statistical dependence among the several criterion variables upsets the significance levels in the series of univariate tests. For details the reader may refer to any of the sources cited under (a) in the first paragraph.

Granting, then, that multivariate analysis should play an even more prominent role in educational research than it presently does, the reader is entitled to ask what sort of multivariate techniques are available. The answer is that practically every univariate statistical technique—from the simplest *t*-test to the most complicated analysis of variance—has its multivariate counterpart. In addition, there are the intrinsically multivariate techniques of discriminant analysis, canonical correlation, factor analysis, etc.

### *Multivariate Analysis of Variance (MANOVA)*

Any time the experimental design is such that an analysis of variance (ANOVA) of some type (one-way classification, multi-factor design with crossed or nested factors, Latin-square design, etc.) would be appropriate if there were but one dependent variable, then MANOVA is applicable when there are two or more dependent variables. To state it briefly: one treats the vector  $[X_1, X_2, \dots, X_p]$  of dependent variables as a single entity, and formulates the null hypotheses in terms of centroids such as  $[\mu_{1r}, \mu_{2r}, \dots, \mu_{pr}]$  (denoting the population centroid corresponding to the *r*th level of the first factor in a two-factor design) instead of means of single variables. The various sums-of-squares (SS) of ANOVA are then

replaced by their corresponding sums-of-squares-and-cross-products (SSCP) matrices, whose diagonal elements are the appropriate SS's of the successive dependent variables, and each off-diagonal element is the appropriate sum-of-cross-products between a pair of dependent variables.

Unlike univariate ANOVA, in which the  $F$ -ratio provides the standard significance test, there are several test criteria available in MANOVA. All of these criteria are functions of the eigenvalues (or latent roots) of the matrix  $S_e^{-1}S_h$ , where  $S_h$  denotes the SSCP matrix for the particular effect being tested (e.g.,  $S_{rc}$  for the interaction effect, corresponding to  $SS_{rc}$  in two-way ANOVA), and  $S_e^{-1}$  is the inverse of the appropriate error SSCP matrix  $S_e$  (e.g.,  $S_w$ , the within-cells SSCP matrix in this case). Letting  $\lambda_1, \lambda_2, \dots, \lambda_p$  be the eigenvalues of  $S_e^{-1}S_h$  in descending order of magnitude (all but the first few of which will ordinarily be zero), the three most widely used test criteria are: (a) Wilks' likelihood-ratio criterion  $[(1 + \lambda_1)(1 + \lambda_2) \dots (1 + \lambda_p)]^{-1}$ , (b) Hotelling's trace criterion  $\lambda_1 + \lambda_2 + \dots + \lambda_p$ , and (c) Roy's largest-root criterion  $\lambda_1/(1 + \lambda_1)$ .

Which of these three criteria is the "best" (in the sense of being most powerful or sensitive in detecting falsity of the null hypothesis) seems still to be a moot question. Schatzoff's (1966) Monte Carlo study suggests that none of them is always the best, since their relative sensitivity depends on the particular way in which the null hypothesis is false. (For example, in the simplest case of one-way MANOVA: Is the centroid of one group "way off" from all the other group centroids, these latter being more or less contiguous; or are all the group centroids about equally disparate from one another?)

In my opinion, the question, "Which criterion is 'best' under what circumstances?" is one that applied research can be of great help in answering since the problem has long defied analytic solution. With the accumulation of studies in which all three criteria are used, we would gain increasing insight into the conditions (i.e., the kind of alternative hypothesis) under which each criterion will be the "best."

### *Discriminant Analysis*

Assuming that a significant effect has been found in MANOVA, the researcher will then want to inquire, in what ways do the groups (or the levels of a given factor) differ from one another? Discriminant analysis provides an answer to this sort of question.

This multivariate technique was invented by Fisher (1936) as a tool for classifying an individual (or a specimen) in one of two groups (species) on the basis of multiple measurements. It was extended to apply to taxonomic problems involving more than two groups by several statisticians, independently but at about the same time (Bryan, 1950; Rao, 1948; Tukey, 1949). It was eventually realized that discriminant analysis could serve



a more basic scientific purpose than taxonomic decisions, by revealing the dimensions along which several groups differ from one another. In this usage, discriminant analysis bears a strong resemblance to factor analysis, except that the latter seeks a parsimonious description of *individual* differences, whereas the former provides a parsimonious description of *group* differences.

When used in conjunction with MANOVA with more than one factor (treatment or classification variable), the notion of "group differences" applies only to the main effects—how do the groups representing the various levels of each factor differ from one another? But discriminant analysis can be used also for studying the nature of interaction effects. In discriminant analysis one determines a linear combination  $Y = v_1X_1 + v_2X_2 + \dots + v_pX_p$  of the original variables so that the relevant  $F$ -ratio of the ANOVA in terms of  $Y$  will be greater than that for any other linear combination. Thus, when the interaction effect is being examined, one would determine the coefficients (or combining weights)  $v_1, v_2, \dots, v_p$  in such a way as to minimize the resulting  $MS_{rc}(Y)/MS_w(Y)$ . This is accomplished by computing the eigenvectors of the matrix  $S_w^{-1}S_{rc}$ , where  $S_w^{-1}$  is the inverse of the within-cells SSCP matrix  $S_w$ , and  $S_{rc}$  is the SSCP matrix for the interaction effect. In general, the optimal linear combination, or the first discriminant function  $Y_1$  as it is called, is determined by using the elements of the eigenvector corresponding to the largest eigenvalue  $\lambda_1$  of the matrix  $S_w^{-1}S_h$  as the combining weights, where  $S_w^{-1}$  and  $S_h$  are as defined earlier. The eigenvectors associated with  $\lambda_2, \lambda_3$ , etc. (through the last non-zero eigenvalue) define the second, third, etc. discriminant functions  $Y_2, Y_3$ , etc. These have the following properties:  $Y_2$  is that linear combination of the  $X_i$ 's which yields the largest  $F$ -ratio  $MS_h(Y_2)/MS_e(Y_2)$  among all linear combinations uncorrelated with  $Y_1$ ;  $Y_3$  yields the largest  $F$ -ratio among those uncorrelated with both  $Y_1$  and  $Y_2$ ; and so forth. It is these linear combinations which define the "dimensions" of group differences or of interaction, and thus facilitate the study of the nature of differences.

This extension of discriminant analysis from a classificatory tool to a more general research technique as a follow-up to MANOVA is probably one of the most significant developments in multivariate analysis during the past ten years. Examples of its use in educational research may be found in Bock (1966), Bock and Haggard (1968), and L. V. Jones (1966). For more detailed discussions of MANOVA and discriminant analysis, the reader is referred to the recent book by Morrison (1967) and the shortly forthcoming ones by Bock (1970) and Tatsuoaka (1970).

It may be pertinent to conclude this section with a caveat: the availability of diverse multivariate techniques should not lull the potential user into thinking that he can "throw everything into the hopper" and come

out with meaningful results. The need for a carefully planned design and a judicious choice of variables still holds—if anything, with even greater force than in univariate analysis. Given a good design and a wise choice of variables, multivariate analysis can be a powerful research tool; without them, it can only generate complicated garbage from simple trash.

## Bibliography

- Abramowitz, M. and Stegun, I. A. (editors). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Washington, D. C.: National Bureau of Standards, 1955.
- Baker, F. B. and Collier, R. O. An Empirical Study into Factors Affecting the *F*-Test Under Permutation for the Randomized Block Design. *Journal of the American Statistical Association* 63: 902-11; 1968.
- Blum, Julius R. and Fattu, Nicholas A. Nonparametric Methods. *Review of Educational Research* 24: 467-87; 1954.
- Bock, R. Darrell. Contributions of Multivariate Experimental Designs to Educational Research. *Handbook of Multivariate Experimental Psychology*. (Edited by R. B. Cattell.) Chicago: Rand McNally, 1966.
- Bock, R. Darrell. *Multivariate Statistical Methods in Behavioral Research*. New York: McGraw-Hill, 1970.
- Bock, R. Darrell and Haggard, Ernest A. The Use of Multivariate Analysis of Variance in Behavioral Research. *Handbook of Measurement and Assessment in Behavioral Sciences*. (Edited by D. K. Whitla.) Reading, Mass.: Addison-Wesley, 1968.
- Box, G. E. P. The Exploration and Exploitation of Response Surfaces: Some General Considerations and Examples. *Biometrics* 10: 16-60; 1954.
- Box, G. E. P. and Hunter, J. S. Multi-factor Experimental Designs for Exploring Response Surfaces. *Annals of Mathematical Statistics* 28: 195-241; 1957.
- Box, G. E. P. and Youle, P. V. The Exploration and Exploitation of Response Surfaces: An Example of the Link Between the Fitted Surface and the Basic Mechanism of the System. *Biometrics* 11: 287-323; 1955.
- Bradley, James V. *Distribution-Free Statistical Tests*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.
- Brown, Rex V. Evaluation of Total Survey Error. *Journal of Marketing Research* 4: 117-27; 1967.
- Bryan, Joseph G. *A Method for the Exact Determination of the Characteristic Equation and Latent Vectors of a Matrix with Applications to the Discriminant Function for More Than Two Groups*. Doctor's thesis. Cambridge, Mass.: Graduate School of Education, Harvard Univ., 1950.
- Chernoff, H. and Savage, I. R. Asymptotic Normality and Efficiency of Certain Non-parametric Test Statistics. *The Annals of Mathematical Statistics* 29: 972-94; 1958.
- Collier, R. O.; Baker, F. B.; and Mandeville, G. K. Tests of Hypothesis in a Repeated Measures Design from a Permutation Viewpoint. *Psychometrika* 32: 15-24; 1967.
- Collier, R. O. et al. Estimates of Test Size for Several Test Procedures Based on Conventional Variance Ratios in the Repeated Measures Design. *Psychometrika* 32: 339-53; 1967.
- Cooley, William W. and Lohnes, Paul R. *Multivariate Procedures for the Behavioral Sciences*. New York: John Wiley and Sons, 1962.
- Cox, D. R. *Planning of Experiments*. New York: John Wiley and Sons, 1957.
- Dunn, Olive J. Multiple Comparisons Using Rank Sums. *Technometrics* 6: 241-52; 1964.
- Eden, T. and Yates, F. On the Validity of Fisher's *Z*-Test when Applied to an Actual Sample of Non-Normal Data. *Journal of Agricultural Science* 23: 6; 1933.
- Elveback, L. and Varma, A. A Selected Bibliography on Simulation and Empirical Sampling. *The Statistical News*, No. 16. New York: N. Y. Chapter of the American Statistical Assoc., 1965.
- Federer, W. T. *Experimental Design*. New York: MacMillan, 1955.
- Feldt, L. S. A Test of the Hypothesis that Cronbach's Alpha on Kuder-Richardson Coefficient Twenty is the Same for Two Tests. *Psychometrika* 34: 363-74; 1969.

- Festinger, Leon. The Significance of Difference Between Means Without Reference to the Frequency Distribution Function. *Psychometrika* 11: 97-105; 1946.
- Festinger, Leon and Katz, Daniel (editors). *Research Methods in the Behavioral Sciences*. New York: Dryden Press, 1953.
- Fisher, Ronald A. The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics* 7: 179-88; 1936.
- Forsythe, G. E. and Householder, A. S. (editors). *Monte Carlo Method*. NBS Applied Mathematics Series 12. Washington, D. C.: U. S. Government Printing Office, 1951.
- Friedman, Milton. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *Journal of the American Statistical Association* 32: 675-701; 1937.
- Gates, C. E. Simulation Study of Estimators for the Line Transect Sampling Method. *Biometrics* 25: 317-29; 1969.
- Glass, Gene V and Maguire, Thomas O. *Analysis of Time-Series Quasi-Experiments*. Cooperative Research Project No. 6-8329. Boulder: Laboratory of Educational Research, Univ. of Colo., 1968.
- Gold, Ruth Z. Tests Auxiliary to  $\chi^2$  Tests in a Markov Chain. *The Annals of Mathematical Statistics* 34: 56-74; 1963.
- Good, Irving J. *Probability and the Weighing of Evidence*. New York: Hafner, 1950.
- Good, Irving J. *The Estimation of Probabilities: An Essay on Modern Bayesian Methods*. Cambridge, Mass.: Mass. Institute of Technology Press, 1965.
- Goodman, Leo A. On Simultaneous Confidence Intervals for Multinomial Proportions. *Technometrics* 7: 247-54; 1965.
- Greenberger, M. Method in Randomness. *Communications of the ACM* 8: 177-79; 1965.
- Greenhouse, S. W. and Geisser, S. On Methods in the Analysis of Profile Data. *Psychometrika* 24: 95-112; 1959.
- Hájek, Jaroslav. *Nonparametric Statistics*. San Francisco: Holden-Day, 1969.
- Hammersley, J. M. and Handscomb, D. C. *Monte Carlo Methods*. New York: John Wiley and Sons, 1964.
- Hammersley, J. M. and Morton, K. W. Poor Man's Monte Carlo. *Journal of The Royal Statistical Society, Series B* 16: 23-38; 1954.
- Hanson, M.; Hurwitz, W.; and Madow, W. *Sample Survey Methods and Theory* Vols. I and II. New York: John Wiley and Sons, 1953.
- Hill, Bruce M. Correlated Errors in the Random Model. *Journal of the American Statistical Association* 62: 1387-1400; 1967.
- Hodges, J. L., Jr. and Lehmann, E. L. The Efficiency of Some Nonparametric Competitors of the t-Test. *The Annals of Mathematical Statistics* 27: 324-35; 1956.
- Hodges, J. L., Jr. and Lehmann, E. L. Rank Methods for Combination of Independent Experiments in Analysis of Variance. *The Annals of Mathematical Statistics* 33: 482-97; 1962.
- Hoeffding, W. Some Recent Developments in Nonparametric Statistics. *Review of the International Statistical Institute* 36: 176-83; 1968.
- Holloway, Lois Nelson and Dunn, Olive J. The Robustness of Hotelling's  $T^2$ . *Journal of the American Statistical Association* 62: 124-36; 1967.
- Hudson, J. D. and Krutchkoff, R. G. A Monte Carlo Investigation of the Size and Power of Tests Employing Satterthwaite's Synthetic Mean Squares. *Biometrika* 55: 1968.
- Hull, T. E. and Dobell, A. R. Random Number Generators. *SIAM Review* 4: 230-54; 1962.
- Hull, T. E. and Dobell, A. R. Mixed Congruential Random Number Generators for Binary Machines. *Journal of the Association for Computing Machinery* 11: 31-40; 1964.
- Jeffreys, Harold. *Theory of Probability*. Third Edition. London: Clarendon Press, 1961.
- Jensen, Donald R.; Beus, G. B.; and Storm, George. Simultaneous Statistical Tests on Categorical Data. *The Journal of Experimental Education* 36: 46-56; 1968.
- Johnson, Palmer O. and Moonan, William J. Recent Developments in Statistical Theory. *Review of Educational Research* 21: 389-414; 1951.
- Jones, Lyle V. Analysis of Variance in Its Multivariate Developments. *Handbook of Multivariate Experimental Psychology*. (Edited by R. B. Cattell.) Chicago: Rand McNally, 1966.



- Kahn, Herman. *Use of Different Monte Carlo Sampling Techniques*. The Rand Corporation P-766. Palo Alto, Calif.: the Corporation, 1955.
- Kao, Edward P. C. A Nonparametric Approach to the 2<sup>3</sup> Factorial Design. *Technometrics* 11: 193-96; 1969.
- Kemphorne, O. *The Design and Analysis of Experiments*. New York: John Wiley and Sons, 1952.
- Klahr, D. A Monte Carlo Investigation of the Statistical Significance of Kruskal's Non-metric Scaling Procedure. *Psychometrika* 34: 319-30; 1969.
- Klotz, Jerome. Small Sample Power and Efficiency for the One Sample Wilcoxon and Normal Scores Test. *The Annals of Mathematical Statistics* 34: 624-32; 1963.
- Koch, Gary G. Some Aspects of the Statistical Analysis of "Split-Plot" Experiments in Completely Randomized Layouts. *Journal of the American Statistical Association* 64: 485-505, 1969.
- Koch, Gary G. and Sen, P. K. Some Aspects of the Statistical Analysis of the Mixed Model. *Biometrics* 24: 27-47; 1968.
- Kraft, Charles H. and van Eeden, Constance. *A Nonparametric Introduction to Statistics*. New York: Macmillan, 1968.
- Kyberg, Henry E. and Smoker, Howard E. *Studies in Subjective Probability*. New York: John Wiley and Sons, 1964.
- Ledley, R. S. *Use of Computers in Biology and Medicine*. New York: McGraw-Hill, 1965.
- Lehmann, E. L. Robust Estimation in Analysis of Variance. *The Annals of Mathematical Statistics* 34: 957-66; 1963a.
- Lehmann, E. L. Asymptotically Nonparametric Inference. *The Annals of Mathematical Statistics* 34: 1494-1506; 1963b.
- Lehmann, E. L. Nonparametric Confidence Intervals for a Shift Parameter. *The Annals of Mathematical Statistics* 34: 1507-12; 1963c.
- MacLaren, M. D. and Marsaglia, G. Uniform Random Number Generators. *Journal of the Association for Computer Machinery* 12: 83-89; 1965.
- Marascuilo, Leonard A. Large-Sample Multiple Comparisons. *Psychological Bulletin* 65: 280-90; 1966.
- Marascuilo, Leonard A. and McSweeney, Maryellen. Nonparametric Post Hoc Comparisons for Trend. *Psychological Bulletin* 67: 401-12; 1967.
- Marsaglia, G. Generating Exponential Random Variables. *The Annals of Mathematical Statistics* 32: 899-900; 1961.
- Marsaglia, G. Generating Discrete Random Variables in a Computer. *Communications of the Association for Computer Machinery* 6: 37, 38; 1963.
- Marsaglia, G. A General Method for Producing Random Variables in a Computer. *Proceedings Fall Joint Computer Conference, American Federation Information Processing Society*. Washington, D. C.: Spartan Books, 1966. Pp. 169-73.
- Marsaglia, G. Pseudo Random Normal Numbers, Answer to Query 27. *Technometrics* 10, No. 2: 1968.
- Marsaglia, G. and Bray T. A. A Convenient Method for Generating Normal Variables. *SIAM Review* 6: 260-64; 1964.
- Marsaglia, G. and MacLaren, M. D. A Fast Procedure for Generating Normal Random Variables. *Communications of the Association for Computer Machinery* 7: 4-10; 1964.
- McDonald, B. J. and Thompson, W. A., Jr. Rank Sum Multiple Comparisons in One- and Two-Way Classifications. *Biometrika* 54: 487-97; 1967.
- Meyer, Donald L. Bayesian Statistics. *Review of Educational Research* 36: 503-16; 1966.
- Meyer, Donald L. and Collier, Raymond O. Bayesian Statistics. *Ninth Annual Phi Delta Kappa Symposium on Educational Research*. Itasca, Ill.: F. E. Peacock, 1970.
- Meyer, H. A. (editor). *Symposium on Monte Carlo Methods*. New York: John Wiley and Sons, 1956.
- Miller, Rupert G., Jr. *Simultaneous Statistical Inference*. New York: McGraw-Hill, 1966.
- Morrison, D. F. *Multivariate Statistical Method*. New York: McGraw-Hill, 1967.
- Moses, Lincoln. Non-Parametric Statistics for Psychological Research. *Psychological Bulletin* 49: 122-43; 1952.
- Naylor, T. H. and Burdick, D. S. Design of Computer Simulation Experiments For Industrial Systems. *Communications of the Association for Computer Machinery*, 1966.



- Naylor, T. H. et al. *Computer Simulation Techniques*. New York: John Wiley and Sons, 1968.
- Nemenyi, Peter B. *Distribution-Free Multiple Comparisons*. Doctor's thesis. Princeton, N. J.: Princeton Univ., 1963.
- Noether, Gottfried E. *Elements of Nonparametric Statistics*. New York: John Wiley and Sons, 1967.
- Noether, Gottfried E. Needed—A New Name. In Letters to the Editor. *The American Statistician* 21: 41; 1967.
- Norris, R. C. and Hjelm, H. F. *An Empirical Investigation of the Effects of Non-Normality upon the Sampling Distribution of the Product Moment Correlation Coefficient*. Cooperative Research Project No. 637. Nashville, Tenn.: George Peabody College for Teachers, 1960.
- Pitz, Gordon F. An Example of Bayesian Hypothesis Testing: The Perception of Rotary Motion in Depth. *Psychological Bulletin* 70: 252-55; 1968.
- Pratt, John W. Robustness of Some Procedures for the Two-Sample Location Problem. *Journal of the American Statistical Association* 59: 655-80, 1964.
- Raiffa, H. and Schlaifer, R. *Applied Statistical Decision Theory*. Boston, Mass: Harvard Univ. Press, 1961.
- Ralston, A. R. and Wilf, H. S. *Mathematical Methods for Digital Computers*. Vol. 1. New York: John Wiley and Sons, 1960.
- Ralston, A. R. and Wilf, H. S. *Mathematical Methods for Digital Computers*. Vol. 2. New York: John Wiley and Sons, 1967.
- Rao, C. Radhakrishna. The Utilization of Multiple Measurements in Problems of Biological Classification. *Journal of the Royal Statistical Society, Series B* 10: 159-93; 1948.
- Rao, J. N. K. and Bayless, D. L. An Empirical Study of the Stabilities of Estimators and Variance Estimators in Unequal Probability Sampling of Two Units per Stratum. *Journal of the American Statistical Association* 64: 540-60; 1969.
- Rosenthal, Irene and Ferguson, Thomas S. An Asymptotically Distribution-Free Multiple Comparison Method with Application to the Problem of  $n$  Rankings of  $m$  Objects. *The British Journal of Mathematical and Statistical Psychology* 18: 243-54; 1965.
- Rulon, Phillip J. Distinctions Between Discriminant and Regression Analysis and a Geometric Interpretation of the Discriminant Function. *Harvard Educational Review* 21: 80-90, Spring 1951.
- Rulon, Phillip J. and Brooks, W. Douglas. On Statistical Tests of Group Differences. *Handbook of Measurement and Assessment in Behavioral Sciences*. (Edited by D. K. Whitla.) Reading, Mass.: Addison-Wesley, 1968.
- Sasaki, Kyohei. *Statistics for Modern Business Decision Making*. Belmont, Calif.: Wadsworth, 1968.
- Savage, I. Richard. *Statistics: Uncertainty and Behavior*. Boston: Houghton-Mifflin, 1968.
- Savage, Leonard J. *The Foundations of Statistics*. New York: John Wiley and Sons, 1954.
- Schatzoff, Martin. Sensitivity Comparisons among Tests of the General Linear Hypothesis. *Journal of the American Statistical Association* 61: 415-35; 1966.
- Schmitt, Samuel A. *Measuring Uncertainty, An Elementary Introduction to Bayesian Statistics*. Reading, Mass.: Addison-Wesley, 1969.
- Sen, P. K. On a Class of Aligned Rank Order Tests in Two-Way Layouts. *The Annals of Mathematical Statistics* 39: 1115-24; 1968.
- Smith, J. *Computer Simulation Models*. New York: Hafner, 1968.
- Steele, R. G. D. A Rank Sum Test for Comparing All Pairs of Treatments. *Technometrics* 2: 197-207; 1960.
- Tate, Merle W. and Clelland, Richard C. *Nonparametric and Shortcut Statistics in the Social, Biological, and Medical Sciences*. Danville, Ill.: Interstate Printers and Publishers, 1957.
- Tatsuoka, Maurice M. *Multivariate Analysis for Educational and Psychological Research*. New York: John Wiley and Sons, 1970.
- Tatsuoka, Maurice M. and Tiedeman, David V. Discriminant Analysis. *Review of Educational Research* 24: 402-20; 1954.
- Tatsuoka, Maurice M. and Tiedeman, David V. Statistics as an Aspect of Scientific



- Method in Research on Teaching. *Handbook of Research on Teaching*. (Edited by N. L. Gage.) Chicago: Rand McNally, 1963.
- Tobach, Ethel et al. A Table for Making Rank Sum Multiple Paired Comparisons. *Technometrics* 9: 561-67; 1967.
- Trotter, H. F. and Tukey, J. W. Conditional Monte Carlo for Normal Samples. *Symposium on Monte Carlo Methods*. (Edited by Herbert A. Meyer.) New York: John Wiley and Sons, 1956.
- Tukey, John W. Dyadic Anova, An Analysis of Variance for Vectors. *Human Biology* 21: 65-110; May 1949.
- Tukey, John W. Conclusions vs. Decisions. *Technometrics* 2: 423-83; 1960.
- Ury, Hans. In Response to Noether's letter, "Needed—A New Name." *The American Statistician* 21: 53; 1967.
- Van Gelder, A. Some New Results in Pseudo-Random Number Generation. *Journal of the Association for Computing Machinery* 14: 785-92; 1967.
- Walker, Helen M. and Lev, Joseph. *Statistical Inference*. New York: Henry Holt, 1953.
- Winer, B. J. *Statistical Principles in Experimental Design*. New York: McGraw-Hill, 1962.
- Wolfowitz, J. Additive Partition Functions and a Class of Statistical Hypotheses. *The Annals of Mathematical Statistics* 13: 247-79; 1942.

## AUTHORS

- RAYMOND O. COLLIER Address: University of Minnesota Title: Professor of Educational Psychology Age: 47 Degrees: B.A., Southern Methodist Univ.; M.Ed., Univ. of Okla.; Ph.D., Univ. of Minn. Specialization: Statistics; Experimental Design.
- ROBERT C. LARSON Address: University of Minnesota Title: Instructor Age: 27 Degrees: B.S., Univ. of Minn.; M.S. Ed., Univ. of Utah Specialization: Statistics; Measurement; Experimental Design.
- LEONARD A. MARASCUILO Address: University of California, Berkeley Title: Associate Professor of Education Age: 38 Degrees: B.A. and M.A., Univ. of Minn.; Ph.D., Univ. of Calif., Berkeley Specialization: Educational Research Methodology.
- MARYELLEN McSWEENEY Address: Michigan State University Title: Assistant Professor of Education Age: 30 Degrees: A.B., Ph.D., Univ. of Calif., Berkeley Specialization: Educational Research Methodology.
- DONALD L. MEYER Address: Syracuse University Title: Associate Professor of education and psychology Age: 37 Degrees: B.S., M.A., and Ph.D., Univ. of Minn. Specialization: Bayesian Statistics and Experimental Design.
- MAURICE M. TATSUOKA Address: University of Illinois Title: Professor of Educational Psychology and of Psychology Age: 48 Degrees: B.S., Nagoya Univ. (Japan); M.A., George Peabody College for Teachers; Ed.D., Harvard Univ. Specialization: Applied Multivariate Analysis.

Available late January, 1970

## *Abstracts* of the 1970 Annual Meeting

March 2-6, Minneapolis

One: Paper Sessions. 600 entries. \$3.50 (paper)

Two: Symposia. 300 entries. \$2.50 (paper)

Combination price for One, Two: \$5.00

Order from AERA, 1126 Sixteenth Street, N.W.,  
Washington, D.C. 20036

Write for advance registration for the Annual Meeting

---

Available Now

## 1969 Directory of American Educational Research Association

Members, Officers, Committees, Publications

Order directly from AERA, Washington

Price to Members: \$2.00. Price to Others: \$10.00